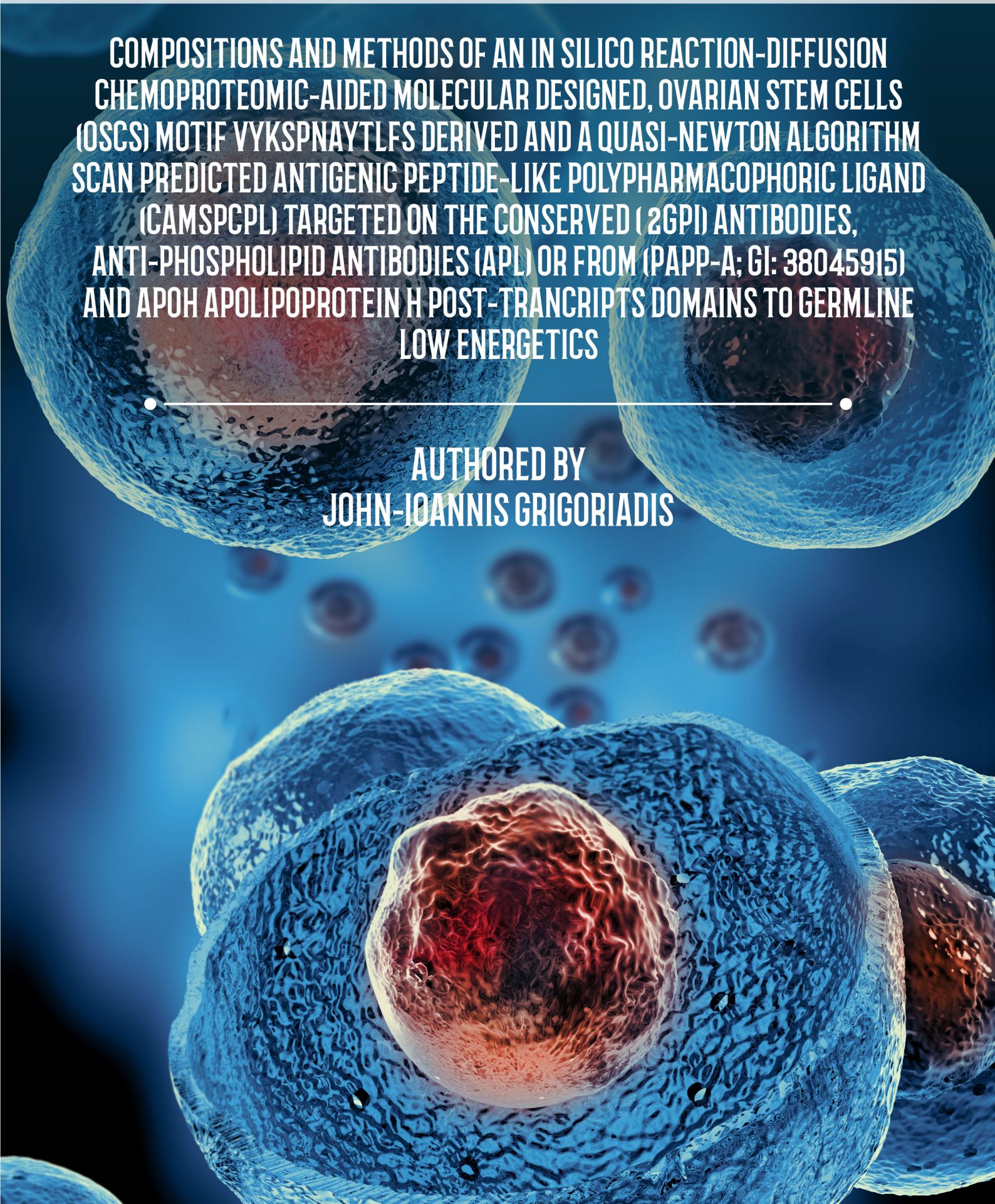
ISBN: 978-1-946628-13-8

COMPOSITIONS AND METHODS OF AN IN SILICO REACTION-DIFFUSION CHEMOPROTEOMIC-AIDED MOLECULAR DESIGNED, OVARIAN STEM CELLS (OSCS) MOTIF VYKSPNAYTLFS DERIVED AND A QUASI-NEWTON ALGORITHM SCAN PREDICTED ANTIGENIC PEPTIDE-LIKE POLYPHARMACOPHORIC LIGAND (CAMSPCPL) TARGETED ON THE CONSERVED (2GPI) ANTIBODIES, ANTI-PHOSPHOLIPID ANTIBODIES (APL) OR FROM (PAPP-A; GI: 38045915) AND APOH APOLIPOPROTEIN H POST-TRANCRIPTS DOMAINS TO GERMLINE LOW ENERGETICS

AUTHORED BY
JOHN-IOANNIS GRIGORIADIS



Crimson Publishers
Wings to the Research

Compositions and Methods of an *in Silico* Reaction-Diffusion Chemoproteomic-Aided Molecular Designed, Ovarian Stem Cells (Oscs) Motif VYKSPNAYTLFS Derived and A Quasi Newton Algorithm Scan Predicted Antigenic Peptide-Like Polypharmacophoric Ligand (CAMSPCPL) Targeted on the Conserved (B2gpi) Antibodies, Anti-Phospholipid Antibodies (apl) or from (PAPP-A; Gi: 38045915) and APOH Apolipoprotein H Post-Trancripts Domains to Germline Low Energetics

John-Ioannis Grigoriadis*

Biogenea Pharmaceuticals Ltd

Greece

***Coressponding author:**

John Ioannis Grigoriadis, Biogenea Pharmaceuticals Ltd.,
Thessaloniki, Greece, PC54623,
Email: biogeneadrug@gmail.com

Publisher By

Crimson Publishers, LLC

May 03, 2018

Content

Abstract

The role of metabolism in ovarian aging is poorly described, despite the fact that ovaries fail earlier than most other organs. Growing interest in ovarian function is being driven by recent evidence that mammalian females routinely generate new oocytes during adult life through the activity of germline stem cells. In this perspective, we overview the female reproductive system as a powerful and clinically relevant model to understand links between aging and metabolism, and we discuss new concepts for how oocytes and their precursor cells might be altered metabolically to sustain or increase ovarian function and fertility in women. Human pregnancy-associated plasma protein-A (PAPP-A) cleaves insulin-like growth factor (IGF) binding protein-4 (IGFBP-4), causing a dramatic reduction in its affinity for IGF-I and II. Through this mechanism, PAPP-A is a regulator of IGF bioactivity in several systems, including the human ovary and the cardiovascular system. PAPP-A belongs to the metzincin superfamily of zinc metalloproteinases, and is the founding member of a fifth metzincin family, the pappalysins. Follicular renewal is also dependent on support of circulating blood mononuclear cells. They induce intermediary stages of meiosis (metaphase I chromosomal duplication and crossover, anaphase, telophase, and cytokinesis) in newly emerging ovarian germ cells, as for the first time demonstrated here, induce formation of granulosa cells, and stimulate follicular growth and development. A pretreatment of OSC culture with mononuclear cells collected from blood of a young healthy fertile woman may cause differentiation of bipotential OSCs into both developing germ and granulosa cells. A small blood volume replacement may enable treatment of ovarian infertility *in vivo*. The transferred mononuclear cells may temporarily rejuvenate virtually all tissues, including improvement of the function of endocrine tissues. Formation of new follicles and their development may be sufficient for IVM and IVF. In this scenario, computer-aided molecular design techniques are offering important tools, mainly as a consequence of the increasing structural information available for these VYKSPNAYTLFS targets by emphasizing the computer-aided molecular design approaches that will be of use to speed-up the efforts to generate new pharmaceutically relevant compounds. Having in mind to represent the signalling of any protein in individual cells whose activity is determined by its post-translational modifications, one can write as many reactions as necessary to sufficiently describe a desired protein dynamics. Michaelis-Menten and Hill kinetics turn protein-protein interactions and transcription of genes into proteins into a system of ODEs. Since the spatial representation of signalling pathways can reveal diffusion patterns that are hidden in ODEs, one might consider a system of PDEs instead of ODEs. Here, we propose for the first time an in silico Reaction-diffusion, Reaction-diffusio, quasi-Newton algorith, Support Vector Machine (SVM), Construction of nonstationary Gabor frames, Discrete finite nonstationary Gabor frames, Numerical complexity, Chemoproteomic-aided Molecular designed, ovarian stem cells (OSCs) motif VYKSPNAYTLFS derived and quasi-Newton algorithm Scan Predicted antigenic peptide-like polypharmacophoric ligand (CAMSPCPL) targeted on the conserved (β2GPI) antibodies, Anti-phospholipid antibodies (aPL) or from (PAPP-A; gi:38045915) and APOH apolipoprotein H post-trancripts domains to Germline low energetics.

Keywords: In silico computer-aided molecular design; Chemoproteomic scan prediction (CAMDCP); Antigenic peptide mimetic; Merged pharmacophore; Anti-beta-2-glycoprotein I (β2GPI) antibodies; Post-trancripts domains; Reaction-diffusion; Quasi-Newton algorithm; Support Vector Machine (SVM); Construction of nonstationary Gabor frames; Discrete finite nonstationary Gabor frames; Numerical complexity; Ovarian stem cells (OSCs); Germline low energetics.

Introduction

Unexplained infertility, sometimes also called idiopathic infertility, refers to the failure to conceive in a couple for whom no definitive cause of infertility can be found. Generally, the duration of infertility is more than 2 years [1,2]. Approximately 10-20% of couples who are unable to conceive are determined to have unexplained infertility. Recurrent abortions are historically defined as three consecutive pregnancy losses before 20 weeks from the last menstrual period. At present, there exist a small number of accepted etiologies for recurrent abortions [3]. Antiphospholipid antibodies (ApLA) are acquired autoantibodies directed to phospholipids that are associated with slow progressive thrombosis and infarction of the placenta. Substances in blood called phospholipids are required for the blood to clot. In some people, the body mistakenly identifies phospholipids as foreign substances and forms antibodies against them. This reaction can be viewed as a confusion of the immune system, called an autoimmune process [4]. APLA represent a family of autoantibodies of different specificities, most of which are directed toward different anionic phospholipids which include cardiolipin, phosphatidylcholine, phosphatidylserine, phosphatidic acid, and phosphatidylethanolamine [5]. Though APLA cause problems, surprisingly 2-15% of the healthy population actually has APLA in their blood. However, these people have very low levels of the antibodies and, therefore, they do not really cause a problem. It is only when they are at high levels that APLA begin to make trouble [6]. APLA are commonly found in people with unexplained infertility, lupus, migraine headache, deep vein thrombosis, and recurrent abortions. Complications caused by APLA include blood clotting, stroke, heart attack, miscarriage, and implantation failure. Anti-phospholipid antibodies: introduction and background Anti-phospholipid antibodies (aPL) comprise a heterogeneous family of autoantibodies that are associated with thrombosis and pregnancy mortality, particularly recurrent pregnancy loss.

There are three clinically relevant aPL:

i) Lupus anticoagulants (LA) that are detected in serum using *in vitro* coagulation assays;

ii) Anti-cardiolipin antibodies (aCL) that are detected in serum or plasma by ELISA and

iii) Anti-beta-2-glycoprotein I (β2GPI) antibodies, which are also detected in serum or plasma by ELISA.

Although ELISA is the most common method of detection of aCL and anti-β2GPI antibodies, other methods such as RIA, fluoro-enzyme immunoassay, multiplexed immunoassay and chemiluminescent immunoassay may also be used to detect aPL positivity (Forastiero et al. 2014). Laboratory criteria for the detection of aPL require detection of LA, aCL or anti-β2GPI on two or more occasions at least 12 weeks apart (Miyakis et al. 2006). It is necessary to conduct repeat tests as some infections are associated with transient false-positive results, especially in aCL tests. Historically, aPL were first identified in patients with biologically false-positive tests for syphilis (Asherson & Cervera 1993). These patients were found to have autoantibodies that produced a false-positive syphilitic 'reagin' test and the antigen in the reagin test is the negatively charged phospholipid, cardiolipin (Pangborn 1941). LA was initially identified in patients with the autoimmune disease systemic lupus erythematosus (SLE). LA were also found to be autoantibodies that reacted with negatively charged phospholipids, and in doing so prolonged the *in vitro* clotting time in phospholipid-dependent coagulation assays, such as the activated partial thromboplastin time and the dilute Russel Viper Venom time. Paradoxically, *in vivo*, LA are associated with thrombotic disease (Love & Santoro 1990, Galarza-Maldonado et al. 2012). Historically, it was noted that several patients with LA also had biologically false-positive tests for syphilis, suggesting that these may be the same autoantibodies. In an attempt to improve the clinical utility and efficiency for these two aPL tests, solid-phase immunnoassays (especially ELISAs) using cardiolipin as the antigen were produced. Some but not all LAs also demonstrate aCL activity and vice versa. Although the traditional view was that aPL bind to negatively charged anionic phospholipids, it is evident that the term 'anti-phospholipid antibodies' is a misnomer, and aPL do not bind directly to phospholipids, but rather to a complex containing a negatively charged phospholipid and a phospholipid-binding 'cofactor' (McIntyre et al. 1997, de Laat et al. 2004, 2007). The most widely studied of these cofactors is β2GPI. Indeed, it has been shown that many, but not all, autoantibodies that are detected in anti-cardiolipin ELISAs can react directly with β2GPI, in the absence of phospholipid, provided that a suitable negatively charged surface is available to immobilize the β2GPI. Immobilization of the β2GPI on such a surface facilitates the binding of aPL in one of two ways: either binding to the negative surface induces a conformational change in the protein that facilitates binding of aPL to a previously hidden cryptic epitope or the negative surface allows clustering of the β2GPI that facilitates high-affinity binding of aPL or a combination of these two mechanisms The relationship between LAs, aCL and anti-β2GPI antibodies varies between patients and specific clones of antibody. Patients may have only one of the three aPL, or any combination of two or three of them, and the relative importance of these three antibodies is not always clear. However, there is growing acceptance that LAs are strongly correlated with disease and that patients who are positive for all three aPL are at the most risk of disease (Ruffatti et al. 2011, Galli 2012, Roggenbuck et al. 2012). LA correlates better with thrombosis, pregnancy morbidity and thrombosis in SLE than do aCL. Detection of LA in a patient with SLE predicts a 50% chance of a thrombotic event over 20 years follow-up. IgG/IgM isotype aCL antibodies, but especially those of the IgG isotype, are frequently associated with thrombosis and pregnancy failure. A recent systematic review indicated that IgG isotype aCL were approximately four times more prevalent than IgM isotype aCL in patients with anti-phospholipid antibody syndrome (APS; Rodriguez-Garcia et al. 2015). This review also suggests that IgA isotype aCL are relatively common in patients with APS (20%) and that IgA isotype anti-β2GPI antibodies were even more common in these patients (56%). IgG and IgM anti-β2GPI are independent risk factors for thrombosis and pregnancy complications (Miyakis et al. 2006).

The APS

With the increased availability of simple tests, it became apparent that aPL identified a subset of patients with SLE who had either

thrombotic disease, pregnancy complications or both. It also became apparent that there were a large number of patients, especially women with pregnancy complications, who had aPL but did not have a defined autoimmune disease or systemic thrombotic disease. Consequently, the APS was defined to describe patients with one or more aPL in the presence of thrombotic disease and/or pregnancy morbidity. The first criteria for classification of APS were developed in Sapporo, Japan, in 1990 and are referred to as the 'Sapporo criteria'. These criteria were revised in 2006 at the Eleventh International Congress of Antiphospholipid Antibodies, and these revised Sapporo criteria remain the most recent international consensus for the diagnosis of APS (Miyakis et al. 2006). Patients can have primary APS in the absence of other disease, or APS secondary to SLE or other pre-existing autoimmune conditions.

Anti-β2GPI antibodies

In recent years in an attempt to improve the diagnostic specificity of aPL testing, there has been a focus on attempting to define specific epitopes within β2GPI with which disease-causing pathogenic antibodies react. Structurally, β2GPI contains five domains of ~60 amino acids that are referred to as short consensus repeats (SCRs) or sushi domains. It has been suggested that the first SCR domain (domain 1) contains an epitope with which particularly pathogenic aPL react (Iverson et al. 1998, de Laat et al. 2009). An international multicentre evaluation of 477 patients established that a strong association existed between anti-β2GPI-domain I antibodies and venous thrombosis, and to a lesser extent pregnancy complications (de Laat et al. 2009). In another recent study, Andreoli et al. evaluated the domain specificity profile of anti-β2GPI antibodies in 159 subjects with persistently positive-, medium- or high-titre anti-β2GPI IgG. The prevalence of anti-β2GPI domain 1 antibodies was 73% in patients with thrombosis, and 64% in patients with obstetric APS. Anti-domain 4 and 5 antibodies were present in 19% and 16% of thrombotic and obstetric APS respectively (Andreoli et al. 2015). Although it is clear that anti-domain 1 anti-bodies are important in the pathogenesis of the systemic thrombotic manifestation of APS, it is much less clear that these antibodies have a particular role in obstetric APS (de Jesus et al. 2014). It is abundantly clear that many women with obstetric as well as other complications of the APS have antibodies that react with other SCR domains of β2GPI (de Jesus et al. 2014, Poulton et al. 2015).

Non-consensus aPL

It is important to acknowledge that other autoantibodies that recognise negatively charged phospholipids or phospholipid-binding proteins, including antibodies against phosphatidylserine, phosphatidylinositol and phosphatidic acid, the zwitterionic phospholipid phosphatidylethanolamine, or the proteins annexin V, protein C, protein S, thrombomodulin and oxidised LDL, have been reported (Gharavi et al. 1987, Oosting et al. 1993, Horkko et al. 1996, Rand et al. 1998, Atsumi et al. 2000, Miyakis et al. 2006, Hirakawa et al. 2012). As phosphatidylserine, such as cardiolipin, is a negatively charged phospholipid, it can also bind to β2GPI, as does oxidised LDL, and it is possible that at least some of the non-consensus aPL are the result of cross-reactive antibodies that will behave as anti-β2GPI or aCL or anti phosphatidylserine (Wu et al. 1999). Although all these antibodies may be related to aPL, they are not included in the consensus criteria for the diagnosis of APS because their clinical significance is unclear (Bertolaccini et al. 2011).

aPL and obstetric diseases

Estimates of the prevalence of aPL in the general obstetric population range from 1 to 7% (Faden et al. 1997); however, most estimates are closer to 1-2%. In contrast, aPL have been reported to be more common in women with stillbirths, recurrent miscarriages, pre-eclampsia and intrauterine growth restriction (IUGR). Approximately 7-25% of unexplained recurrent miscarriage may be due to the presence of aPL (Drakeley et al. 1998) along with 3.8% of stillbirths and 13.7% of IUGR (Silver et al. 2013, Cervera et al. 2015). The reported occurrence of aPL in pre-eclampsia is highly variable, ranging from 11% to 61% (Branch et al. 1989, Yasuda et al. 1995, Katano et al. 1996); however, one group has shown that aPL are one of the strongest maternal risk factors for developing pre-eclampsia, increasing a woman's risk approximately ten-fold (Duckitt & Harrington 2005).

The revised Sapporo criteria define the obstetric manifestations of APS as follows:

i) The occurrence of recurrent spontaneous abortion (miscarriage) before the tenth week of gestation,

ii) Unexplained fetal loss after the tenth week of gestation or

iii) The development of pre-eclampsia/eclampsia or placental insufficiency before the 34th week of gestation (Miyakis et al. 2006). Recently, the Taskforce on Obstetric Antiphospholipid Syndrome has systematically reviewed the relevant literature and confirmed the association of aPL with these complications of pregnancy (de Jesus et al. 2014). The first demonstration that aPL are directly pathogenic in pregnancy came in 1990 with the passive transfer of human aPL into pregnant mice (Rote et al. 1990). This resulted in fetal demise, and since then several groups, using either passive transfer of aPL or immunization with β2GPI, have confirmed that aPL cause fetal demise and growth restriction equivalent to stillbirth/miscarriage and placental insufficiency using murine models (Bakimer et al. 1992, Blank et al. 1994). Moreover, a role for β2GPI has also been confirmed in the development of aPL-mediated pregnancy loss (Robertson et al. 2004). Until recently, whether aPL cause pre-eclampsia, a human pregnancy-specific disease characterized by new-onset hypertension in the second half of pregnancy, had not been demonstrated. However, when pregnant mice are immunized with β2GPI, aPL titres increase and the mice exhibit pre-eclampsia-like symptoms such as hypertension and proteinuria (Ding et al. 2014). How aPL cause the pregnancy complications with which they are associated remains unclear. In both women and men, aPL are associated with systemic thrombotic disease, and that aPL can also induce thrombosis has also been demonstrated in animal models where transfer of human aPL to mice results in excess thrombosis, but usually only following an induced vascular insult. It is now very

clear that thrombosis is not the primary mechanism by which aPL induce obstetric complications (Viall & Chamley 2015); although thrombosis may be seen in a small proportion of placentae from late fetal deaths, uteroplacental thrombosis is not a feature of early gestation losses (Viall & Chamley 2015).

Before the widespread access to the ELISA-type aPL assays, women with these antibodies were usually identified because of a poor obstetric history, specifically late fetal death, which was frequently accompanied by placentae with evidence of extensive infarction (Out et al. 1991). Many, if not most of these, women also had classical autoimmune diseases and systemic thrombotic complications. This finding, coupled with the known association of aPL with thrombotic disease, leads to the reasonable hypothesis that aPL-mediated fetal demise was due to thrombosis in the uterine spiral arteries that supply maternal blood to the placenta, or in the intervillous spaces of the placenta (the original basis for the use of anticoagulant therapies such as heparin and low-dose aspirin (LDA; Ginsberg et al. 1995). However, with easy access to tests for aPL, a large number of women have been identified with aPL who do not have systemic thrombotic disease and whose placentae were not characterized by thrombotic lesions/infarction (Out et al. 1991, Salafia & Cowchock 1997). This led to the suggestion that obstetric manifestations of APS are not mediated primarily by thrombosis, but rather by direct effects of aPL on the trophoblast cells of the placenta. This was further supported by the finding that β2GPI is synthesised by the trophoblast and is thus endogenously localized to the syncytiotrophoblast, cytotrophoblast and extravillous trophoblast populations (Chamley 1997). Indeed, a recent systematic review concluded that only one-third of women with aPL demonstrate signs of uteroplacental thrombosis (Viall & Chamley 2015).

There is no causative relationship between infertility and aPL

The incidence of aPL is reported to be 22 and 30%, respectively, in infertile women and women with recurrent IVF failure (Buckingham & Chamley 2009). However, due to a lack of well-designed studies, and lack of concordance between antibodies detected in each study, the causative nature of this association has not been established. A systematic review of the literature analysed the association of aPL and IVF outcome and showed that 13 of 29 studies reported a higher prevalence of aPL in infertile women. However, most of the studies used non-standardized, non-criteria aPL tests. Importantly, the presence of aPL was not found to influence IVF outcome, and treatment of aPL-positive patients undergoing IVF was not found to be beneficial. The Taskforce on Obstetric Antiphospholipid Syndrome also recently reviewed this topic and concluded that there is little evidence that aPL affect negatively on fertility and that infertility should not be included in the criteria for APS (de Jesus et al. 2014).

How do aPL cause pregnancy complications?

aPL induce characteristic histologic lesions in the placenta

If we are to understand how aPL induce pregnancy complications via direct effects on placental trophoblasts, it might be expected that we would have a clearer understanding of the pathologic lesions these antibodies induce in the placenta of affected pregnancies. In the past 30 or so years, there have been a number of small studies reporting histopathologic lesions in placentae from women with aPL; however, there has been no overview of these studies. A recent systematic review of the literature clarified which lesions are associated with aPL and reported that there is a 'fingerprint' of five common pathologic features in the placentae of women with aPL (Viall & Chamley 2015). These five lesions are as follows:

i) Placental infarction (found only in later gestation placentae)

ii) Impaired remodelling of uterine spiral arteries

iii) Decidual inflammation

iv) Increased syncytial knots in the syncytiotrophoblast and

v) Decreased vasculosyncytial membranes (Viall & Chamley 2015). Thrombotic lesions are not common in the placentae of women with aPL and were strikingly absent from placentae of early gestational losses, confirming that thrombosis is not a primary pathophysiological mechanism in these pregnancies (Salafia & Cowchock 1997).

Beginning with the experiments that revealed the microscopic mechanisms of nerve cell excitation [7], measurements of ion currents through nanoscale channels and pores have become the basis of many experimental techniques in biology and biotechnology. In addition to permitting the study of the behavior of individual proteins that allow the passage of ions into and out of cells [8], ion current measurements through nanopores have been used to study the rupture of molecular bonds [9-11], to distinguish between similar molecules [12], and to determine the properties and sequences of nucleic acid molecules [13-17]. However, since direct experimental imaging of molecules within nano-pores is extremely difficult, computation plays an important role in associating current with nanoscale phenomenon [18-25] (see [26,27] for recent reviews of the field). Equilibrium and transport properties of ionic solutions can be simulated using explicit ion methods such as all-atom molecular dynamics [16,20] or Brownian dynamics [28-30], or by using continuum models such as the Poisson-Boltzmann and Poisson-Nernst-Planck equations [31,32]. While the explicit ion methods provide the most accurate description of the system's behavior, both in spatial and temporal domains, they are stochastic in nature and thus require long, computationally expensive simulations to obtain average properties. Furthermore, the application of an explicit ion method usually requires the system to be described with the same resolution over the entire simulation domain. Often, this leads to a situation where a majority of the computational effort is applied to simulate a nearly uniform solution where

quantities of interest exhibit little variation. In contrast, continuum methods allow different regions of the same system to be described at varying levels of detail, and thus focus the computational effort on regions that require a more precise description. In addition to being more computationally efficient, continuum models more easily incorporate certain types of boundary conditions that arise in physical systems, such as boundaries of fixed concentration or electrostatic potential.

The traditional continuum approach to modeling ionic transport is based on the Poisson-Nernst-Planck equations (PNPE). Although the PNPE have been applied successfully to model the electro-diffusion phenomena [33,34] the equations are not without drawbacks. Within the PNPE approach, ions are modeled as mathematical points of negligible physical dimension, thereby allowing for accumulation of ions at unrealistically high concentrations in certain regions of the system. A modified formulation of the PNPE, called the modified Poisson-Nernst-Planck equations (MPNPE) [35], explicitly takes the physical dimensions of ions into consideration, which limits the maximum concentration that attained in the system. The advantage of using MPNPE over PNPE becomes apparent in the systems that contain regions subject to strong attractive potentials, for example, near charged surfaces. The finite difference method has been widely used to solve the Nernst-Planck equations in one or three dimensions [33,36-38]. Although the finite difference method is straightforward to implement, applying this method to systems that have curved boundaries and complicated geometries is challenging. In this respect, using a finite element method is more appropriate as it naturally handles complex geometries, such as the molecular surfaces of DNA molecules and ion channels. Finite element methods for solving the three-dimensional PNPE have already been described [39,40]. However, numerical studies of the MPNPE have been limited to one-dimensional systems [34] or the three dimensional spherical case [41] and have not been applied to simulate ion flow through a solid-state nanopore, which is the main process considered in this work. In contrast to the previous efforts, our finite element method conserves ion concentration, takes into account the sharp repulsive potentials present near the walls of an ion channel [19,33], and is able to reproduce the results of explicit ions simulations. The presence of a sharp, repulsive potential at the interface of fluid and solid-state domains necessitates formulation of a new stable numerical method for finding the solution of the PNPE and the MPNPE.

Structure of the human placenta and the role of trophoblasts

It is now widely accepted that aPL act directly upon trophoblasts of the placenta to induce pregnancy complications in women, and indeed, these antibodies can be eluted from affected placentae (Chamley et al. 1993). The anatomy of the human materno-fetal interface is quite unique and differs significantly from that in commonly used laboratory animals such as rodents [42]. The human placenta has a villous or tree-like branching structure. The body of the human placenta is covered on its maternal-facing aspect by a single multinucleated cell, the syncytiotrophoblast, which is bathed in maternal blood and has a surface area of 11-13 m2 at term (Mayhew 2008). Therefore, the syncytiotrophoblast acts as both a barrier between the maternal and fetal organisms and also the link between the two [43]. The syncytiotrophoblast produces vast quantities of hormones and other factors that are crucial to the success of pregnancy; yet, it is not mitotically active and is formed and replenished by fusion of underlying villous cytotrophoblasts [44]. Growing out from points of contact between the maternal uterine decidua and the villous placenta are columns of extravillous trophoblasts. The extravillous trophoblasts most proximal to the placenta can proliferate, but as the cells migrate away from the villous surface, they lose their ability to proliferate and instead gain an invasive phenotype. These extravillous trophoblasts migrate into the decidua where they invade and transform the uterine spiral arteries that supply maternal blood to the placenta. The extravillous trophoblasts erode the musculo-elastic walls of the spiral arteries such that these vessels are transformed from narrow-bore resistance arteries to large-diameter tubes that lack the capacity to respond to vaso-constricting stimuli. This transformation of the spiral arteries is necessary to allow the large and uninterrupted supply of maternal blood into the placenta that is required for increasing fetal growth in the later part of pregnancy [45]. As a terminal phase of differentiation, at least some of the extravillous trophoblasts from early gestation form multinucleated placental bed giant cells, the function of which is unknown, but are likely to be important sources of hormones (Pijnenborg et al.). aPL have been reported to affect the function of each of these populations of trophoblasts *in vitro*.

In vitro effects of aPL on trophoblasts

Several *in vitro* studies have been conducted to determine the mechanisms by which aPL interact with and directly affect trophoblast function. These studies were systematically reviewed recently by Tong et al. (2015). That systematic review showed that there is strong *in vitro* evidence that aPL detrimentally affect trophoblast syncytialization, viability and invasion, and this *in vitro* data support the lesions seen in the placentae of aPL-affected pregnancies (Viall & Chamley 2015). Furthermore, aPL also disrupt molecular signalling mechanisms in trophoblasts and increase inflammation and coagulation on the cell surface.

Trophoblast proliferation, migration and invasion

aPL may negatively affect trophoblast viability and function by reducing proliferation and invasiveness. The majority of studies have examined the effects of polyclonal aPL isolated from patient sera, as well as murine monoclonal aPL on placental explants or cell lines (Yacobi et al. 2002, Ornoy et al. 2003, Bose et al. 2004, Schwartz et al. 2007). Studies examining the mechanism of reduced migration showed that aPL mediate this effect possibly through the down-regulation of interleukin 6 (IL6; Mulla et al. 2010) and the up-reguation of TIMP2 (Albert et al. 2014), all independently of Toll-like receptor 4 (TLR4). Another study, however, has implicated TLR4 in reduced trophoblast migration by aPL (Poulton et al. 2015). More recently, apolipoprotein E receptor 2 (ApoER2) has been implicated in aPL-reduced trophoblast migration (Ulrich et al. 2016). Others, however, have suggested that aPL binding directly to trophoblasts via adhered β2GPI reduce invasion by altering the repertoire of cell adhesion molecules (Di Simone et al. 2000a, 2002). A consequence of this reduced migratory response has been demonstrated using an *in vitro* matrigel model of spiral artery transformation. In this system,

aPL and sera from APS patients with pregnancy morbidity disrupt the normal trophoblast-endothelial cell interactions (Alvarez et al. 2015).

aPL have also been shown to inhibit the proliferation of trophoblasts and the final differentiation of extravillous trophoblasts into giant multinucleated cells *in vitro* (Quenby et al. 2005). Reduced proliferation of cytotrophoblasts in response to aPL would have at least two consequences.

i) As villous cytotrophoblasts are the proliferating progenitors of the syncytiotrophoblast, reducing the proliferation of villous trophoblasts may limit the growth and repair of the syncytiotrophoblast, resulting in reduced placental transport and barrier functions.

ii) Proliferating cytotrophoblasts are also progenitors for extravillous trophoblasts, and thus, reduced trophoblast proliferation would lead to a limited pool of extravillous trophoblasts, resulting in inadequate invasion of the spiral arteries, underperfusion, oxidative stress and ischaemia-reperfusion injury in the placenta. These, in turn, could contribute to the early pregnancy loss, pre-eclampsia and IUGR that are often seen in women with aPL (Tong et al. 2015). The same consequences would result from an aPL-induced reduction in invasiveness of extravillous trophoblasts.

Trophoblast death

Several *in vitro* studies have reported an increase in trophoblast death in response to aPL using either term trophoblasts or first-trimester trophoblasts (Di Simone et al. 2001, Yacobi et al. 2002, Ornoy et al. 2003, Schwartz et al. 2007, Chen et al. 2009, Mulla et al. 2009). Cytotrophoblasts isolated from term placentae (Di Simone et al. 2001, 2006) have been shown to alter their expression of the apoptotic regulators Bax and Bcl2 in response to aPL, but without overt signs of cell death (Di Simone et al. 2006). The final stage in the normal life cycle of the multinucleated syncytiotrophoblast is thought to be that in localised regions, programmed cell death is initiated, followed by the extrusion of proapoptotic or apoptotic nuclei in large multinucleated vesicles called syncytial nuclear aggregates (SNAs; Mayhew 2008, Askelund & Chamley 2011). These SNAs are then deported via the blood to the maternal lungs where they become trapped in small vessels and are cleared, most likely, by pulmonary endothelial cells. There is evidence that aPL alter the nature of cell death in the syncytiotrophoblast to a more necrotic process, evidenced by reduced expression of executioner caspases 3 and 7 (Chen et al. 2012), as well as reduced expression of the protective apoptotic regulator TRAIL in placental explants in response to aPL (Pantham et al. 2012). We have recently demonstrated that aPL are internalized into the syncytiotrophoblast of both first-trimester and term placentae via an antigen-specific receptor-mediated process (Viall et al. 2013). Once inside the syncytiotrophoblast, aPL bind to mitochondria and disrupt mitochondrial function and also induce the release of proapoptotic cytochrome c from the mitochondria (Viall et al. 2013, Pantham et al. 2015a). Consequently, there is an increase in the number of SNAs extruded from the syncytiotrophoblast (Chen et al. 2009, Pantham et al. 2015a). The proteome of SNAs from aPL-treated placentae is altered compared with that of SNAs from control antibody-treated placentae with notable changes in mitochondria-related proteins (Pantham et al. 2015a). In contrast to the SNAs from normal placentae, which prevent the activation of endothelial cells, the SNAs released from aPL-treated first-trimester placentae activate endothelial cells, with the likelihood that the changes in both the number and nature of SNAs contributes to the maternal endothelial cell activation, which is a hallmark of pre-eclampsia (Viall et al. 2013). Metabolomic analysis of the culture media from first-trimester placental explants treated with aPL has demonstrated an increase in ceramide metabolites, which play key roles in cell death by suppressing protein kinase C-epsilon, which is also found to be reduced in response to aPL (Pantham et al. 2015b). It appears likely that aPL alter cell death processes in many if not all trophoblast populations in the human placenta. Increased death in the syncytiotrophoblast is likely to contribute to increased production of SNAs, whereas increased death of extravillous trophoblasts is likely to reduce the pool of these cells available to modify the spiral arteries (Quenby et al. 2005).

Cytokine and hormone production

The deciduae of women with APS show signs of inflammation (Tong et al. 2015). *In vitro* studies suggest that aPL may alter the cytokine and hormonal milieu produced by trophoblasts, leading to a pro-inflammatory environment at the maternal-placental interface and altered trophoblast function (Abrahams 2009). It was shown nearly two decades ago that aPL reduce the basal and GnRH-induced production of human chorionic gonadotropin (hCG; Katsuragawaet al. 1997, Di Simone et al. 2000b, Schwartz et al. 2007). Recent evidence suggests that this aPL-mediated reduction in hCG secretion by trophoblasts is mediated via TLR4 (Marchetti et al. 2014). Production of hCG is usually considered to be a marker of syncytiotrophoblast endocrine function, as this cell is the major source of hCG; however, it has been shown that hyperglycosylated hCG can be secreted by, and promotes, extravillous trophoblast invasion (Fournier et al. 2015). Thus, the reductions in hCG secretion induced by aPL may reflect adverse effects of the antibodies on both syncytiotrophoblast and extravillous trophoblast functions. Production of IL3 may be reduced in women with aPL-mediated fetal loss and IL3 has been shown to overcome the inhibitory effects of aPL on trophoblast invasion and hCG secretion, but not trophoblast proliferation (Di Simone et al. 2000a, Chamley et al. 2001). *In vitro* studies utilizing first-trimester trophoblasts and murine monoclonal aPL have demonstrated that aPL (both murine mAb and patient-derived aPL) increase production of IL8, IL1β, monocyte chemotactic protein 1 (MCP1) and GROα in a TLR4-dependent manner [46]. Furthermore, downstream of TLR4, IL1β secretion is mediated by the induction of endogenous uric acid, which in turn activates the Nod-like receptor, Nalp3, leading to Nalp3/ASC/caspase 1 inflammasome activation and subsequent IL1β processing and secretion (Mulla et al. 2013a). These aPL-induced changes could contribute to the reduced placental invasion and pro-inflammatory profile reported in pregnancies affected by aPL (Abrahams 2009). Reported changes in trophoblast/placental cytokine production induced by aPL are summarized in (Table 1).

Coagulation

There has been a considerable amount of research into the mechanisms of coagulation and thrombosis leading to placental infarction due to aPL. aPL may cause coagulation in the proximity of the placenta by reducing the expression of annexin A5 on placental villi (Rand et al. 2006). Annexin A5, also known as placental anticoagulant protein-I and vascular anticoagulant-α, is thought to function as a potent anticoagulant by competing with coagulation factors for binding sites on anionic phospholipids such as phosphatidylserine [47]. Annexin A5 is constitutively expressed on the apical surface of syncytiotrophoblast and may be necessary to maintain vascular homeostasis and blood fluidity, functioning as an 'anticoagulant shield' in the placenta during pregnancy. Disruption of this shield due to aPL has been observed in placental villi from women with APS, as well as in trophoblasts cultured with aPL *in vitro* (Krikun et al. 1994, Rand & Wu 1999, Rand et al. 1994, 1997, 1998, 2003, 2004, 2006). It has therefore been postulated that the disruption of the annexin A5 anticoagulant shield in the placenta results in accelerated coagulation and thrombosis at the placental surface, causing recurrent fetal loss and fetal growth restriction in pregnancies with aPL, and this is supported by data from an annexin A5 knockout mouse model (Rand et al. 1994, 2006, Ueki et al. 2012). However, the disruption of the annexin A5 shield is not a universal feature of placentae from pregnancies complicated by aPL (Lakasing et al. 1999). Furthermore, enhanced coagulation at the placental surface is not a common histopathologic feature of pregnancies affected by aPL (Tong et al. 2015), and thus, it is unclear how important disruption of the annexin A5 anticoagulant shield is in the pathogenesis of obstetric APS.

Receptors that mediate the effects of aPL in the placenta

Although it was originally thought likely that aPL reacted with phospholipids in the cell membrane of target cells, it now seems much more likely that the interaction of aPL with cells is predominantly via cell surface receptors. Several receptors have been implicated in aPL signalling including TLRs 2 and 4, which have lipid-based ligands, as well as lipoprotein receptors with which β2GPI interacts [48-50]. The ability of aPL to activate these TLRs may arise because β2GPI shares molecular mimicry with bacterial products such as lipopolysaccharide (Blank & Shoenfeld 2004, Sorice et al. 2007). Interactions of aPL with endothelial cells have been shown to be mediated by TLR4, and when they have invaded the spiral arteries, extravillous trophoblasts take on the functions and some feature of endothelial cells. Although the effects of aPL on extravillous trophoblasts have been shown to be mediated by TLR4 (Mulla et al. 2009), there is some disagreement as to whether the intracellular signalling is mediated primarily by the MyD88 or TRIM/TRAF pathways (Poulton et al. 2015). It has also recently been shown that aPL are rapidly taken up specifically into the syncytiotrophoblast by a receptor-mediated mechanism. The receptor responsible for this uptake has not yet been specifically identified; however, it was shown that this was not typical antibody transport and was not mediated by an Fc-receptor [51]. The receptor appeared to be a member of the LDL receptor family, as its effects were blocked by receptor-associated protein (Viall et al. 2013). Indeed, in other work, the LDL receptor family member ApoER2 has been shown to mediate changes in trophoblast function and pregnancy complications induced by aPL (Ulrich et al. 2016).

Energy, aging and the role of mitochondria

Early in life's history, a complex signaling network evolved to maximize the number of descendants a cell could produce in a particular environment. This ancient network, which still exists in cells today, promotes growth and reproduction when the environment is favorable and suppresses these activities during harsh times (Kirkwood, 1987). This system explains in large part why many species gain health benefits from dietary restriction (DR) and how the body adapts to changing supplies and demands for energy. As we learn more about this survival network, it is becoming increasingly plausible to stimulate it pharmacologically. Indeed, molecules that mimic DR are in development for treatment of many aging-related health issues, such as type II diabetes, inflammation and muscle degeneration (Blum et al., 2011; Chiba et al., 2010). Despite rapid progress in this area, one aspect of human health that has been largely neglected is reproductive potential. New molecular modeling approaches, driven by rapidly improving computational platforms, have allowed many success stories for the use of computer-assisted drug design in the discovery of new mechanism-or structure-based drugs [52]. In this overview, we highlight three aspects of the use of molecular docking whether one is working on metabolism or aging; it is hard to ignore mitochondria, the structures at the center of cellular energy production and utilization. Early in life's history, a complex signaling network evolved to maximize the number of descendants a cell could produce in a particular environment. This ancient network, which still exists in cells today, promotes growth and reproduction when the environment is favorable and suppresses these activities during harsh times (Kirkwood, 1987). This system explains in large part why many species gain health benefits from dietary restriction (DR) and how the body adapts to changing supplies and demands for energy [53]. As we learn more about this survival network, it is becoming increasingly plausible to stimulate it pharmacologically. Indeed, molecules that mimic DR are in development for treatment of many aging-related health issues, such as type II diabetes, inflammation and muscle degeneration (Blum et al., 2011; Chiba et al., 2010). Despite rapid progress in this area, one aspect of human health that has been largely neglected is reproductive potential [54,55]. New molecular modeling approaches, driven by rapidly improving computational platforms, have allowed many success stories for the use of computer-assisted drug design in the discovery of new mechanism-or structure-based drugs. In this overview, we highlight three aspects of the use of molecular docking [56-59].These organelles are essential for generating most of the ATP in the body, which in humans amounts to about 65 kg per day to meet basic metabolic demands (Tornroth-Horsefield and Neutze, 2008). Other key functions of mitochondria include calcium buffering, reduction-oxidation (redox) homeostasis and programmed cell death (apoptosis) [60,61]. Mitochondria are continuously moving throughout the cell-undergoing fusion, fission, and degradation-to eliminate and replace damaged organelles, and to meet fluctuating energy needs (Palmer et al., 2011). Data from flies, rats, mice, monkeys and humans show that as tissues age, both the number and activity of mitochondria decline, compensated by an increase in

their overall size (Cho et al., 2011; Ferguson et al., 2005; Short et al., 2005; Wallace, 2001). Mitochondrial dysfunction is associated with, and potentially contributes to, common aging-related diseases such as atherosclerosis, obesity-induced type II diabetes, sarcopenia and neurodegenerative disorders (Di Lisa et al., 2009; Lin and Beal, 2006; Wallace, 2001). Underlying processes include a decline in mitochondrial membrane potential and ATP output, increased activation of the mitochondrial permeability transition pore (mPTP), mitochondrial membrane depolarization and leakage of mitochondrial matrix solutes into the cytoplasm (Di Lisa et al., 2001; Hafner et al., 2010; Liu et al., 2011; Wallace, 2001). In addition to aging and aging-related diseases, considerable evidence supports a role for mitochondria as mediators of the benefits of DR in rodents and humans (Cerqueira et al., 2011; Civitarese et al., 2007; Lopez-Lluch et al., 2006; Lopez-Lluch et al., 2008). In a variety of species (Guarente, 2008; Johannsen and Ravussin, 2009), DR increases mitochondrial number and function [62,63]. A recent rat study, however, did not observe this change (Hancock et al., 2011). The sirtuin-1 (SIRT1)-AMPK network, considered a mediator of DR physiology, acts to raise both the number and activity of mitochondria (Gerhart-Hines et al., 2007), as do DR mimetics that stimulate SIRT1 or AMPK activity, such as resveratrol (Baur et al., 2006; Feige et al., 2008; Funk et al., 2010; Lagouge et al., 2006), SRT1720 (Minor et al., 2011) and metformin (Canto et al., 2009; Suwa et al., 2006). In flies and nematodes, changes in mitochondrial metabolism are known to be necessary (Bahadorani et al., 2010; Bishop and Guarente, 2007; Zid et al., 2009) and sufficient (Bahadorani et al., 2010; Durieux et al., 2011; Rera et al., 2011) for DR to extend lifespan [64,65].

In this work we provide the systems and the governing equations. Our nonlinear finite element method for solving the MPNPE is also described with the results of several computational experiments that highlight the utility of the MPNPE and the necessity of having a stabilized algorithm [66-70]. Also in this section, we provide the Galerkin formulation for the equations that do not have large drift terms, and a streamline-upwind-Petrov-Galerkin (SUPG) method for the equations in which such terms are present [71-74]. We describe from a mathematical point of view, post-translational modifications can be advantageously modelled as enzyme reactions where a modifying protein (providing a phosphate group or ubiquitin) is considered to act as enzyme, i.e. it binds a substrate to form a compound, modifies the substrate in the compound and is released from the compound, becoming available to modify other target substrates [75,76]. Mathematically, enzyme reactions can be described by using the law of mass action and the quasi-steady-state approximation yielding thus a non-linear term for a gain of the modified protein (the product of the reaction) and a loss of the former state of the protein (the substrate entering the reaction) [67,68,77]. Over the last few decades, we have seen the formulation of group iterative methods for solving the two dimensional elliptic partial difference Equations [78-81]. In 1991, a half-sweep iterative method had been introduced by Abdullah [82] via the explicit decoupled group (EDG) iterative method which was shown to be faster and computationally economical than the existing explicit group (EG) method due to Yousif and Evans [81] for solving elliptic partial differential equation. Inspired by Abdullah [82], Othman and Abdullah [83] proposed quarter-sweep iteration through the modified explicit group (MEG) method. Following this, Ali and Ng [84] extended the idea and formulated the modified explicit decoupled group (MEDG) method for solving two-dimensional Poisson equation. The MEDG method exhibits a better convergence rather than the existing group schemes of the same family, namely EG, EDG and MEG methods. Over the last few decades, cubic splines have been widely used to approximate differential equations due to their ability to produce highly accurate solutions. In this paper, the numerical solution of a two-dimensional elliptic partial differential equation is treated by a specific cubic spline approximation in the x-direction and finite difference in the y-direction. A four point explicit group (EG) iterative scheme with an acceleration tool is then applied to the obtained system. The formulation and implementation of the method for solving physical problems are presented in detail. The complexity of computational is also discussed and the comparative results are tabulated to illustrate the efficiency of the proposed method. In 1986, Yousif and Evans [81] developed the explicit group (EG) iterative method where a small group of 2, 4, 9, 16 and 25 points were constructed in the iterative processes for solving Laplace's equation.This approach turns enzyme reaction kinetics into ordinary differential equations (ODEs) broadly used by mathematicians. One of the simplest 'protein-mRNA' models is a non-linear biochemical oscillator proposed by Goodwin in 1965 [69] that simulates expression of a single gene controlled by its protein product P, thus, closing a negative feedback loop. In particular, equations appearing in the original Goodwin model are, respectively,

$$\text{dm RNAdt} = a1A1+k1P-\delta1,\ \text{dPdt}=a2\text{mRNA}-\delta2$$

reporting here only production (transcription in the first equation and translation in the second) and degradation terms (δ1 and δ2).Based on this model, a compartmental ODE model for the circadian clock rhythm of the (β2GPI) antibodies, Anti-phospholipid antibodies (aPL) protein expressed during a day in Infertility was proposed by Leloup and Goldbeter [70]. The model includes expression of the (β2GPI) antibodies, Anti-phospholipid antibodies (aPL) gene in the nucleus, a process that is influenced by available light, and translation of (β2GPI) antibodies, Anti-phospholipid antibodies (aPL) mRNA into (β2GPI) antibodies, Anti-phospholipid antibodies (aPL) in the cytoplasm that, afterwards, enters the nucleus and negatively regulates the transcription of its gene, as it is observed *in vivo* in [71]. More complicated 'protein-protein' and 'protein-mRNA' ODE models have been designed to represent intracellular signalling of the protein apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b [72-81]. One of them is also proposed in [82] and it simulates initial activation and regulation of apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b in response to DNA damage. Activation is performed through the interactions with the ATM protein, that is identified as DNA double strand break (DSB) sensor [83], and regulation of apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b by the proteins Wip1 and Mdm2, on whose genes apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b act as a transcription factor [84,85]. These four proteins have been identified as necessary to create a minimal apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b network yielding sustained oscillations in apolipoprotein H [Homo sapiens]

and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b concentration *in vivo* [72,86]. The spatial organisation of the cell, however, suggests to consider migration of the involved species in the cytosol and between the compartments, and thus to model a protein signalling more realistically by including diffusivities of the species and by setting particular translocation conditions to model exchanges of the species between the compartments. Thus, one has to deal with coupled systems of nonlinear evolution partial differential equations (PDEs) for the concentrations of the species in the nucleus and in the cytoplasm with transmission boundary conditions (BCs) imposed on the inner nuclear envelope and on the outer cellular membrane (Robin-like and zero-flux BCs in our modelling setting). Among other mathematical methods that are available in the literature, the semi-implicit Rothe method [87] can be used when dealing with non-linear systems of evolution equations. Without going into details, let us mention that after a suitable linearisation of non-linear equations, one can discretise time derivatives and solve a finite number of elliptic equations for which a non-negative and unique solution is guaranteed by the Lax-Milgram theorem [87,88]. A Rothe function formed from the solutions of the elliptic problems then can serve as a good approximation of a solution of the original reaction-diffusion problem (that is also non-negative and unique) in an entire time interval for the problem under consideration. The aim of this article is to acquaint readers with reaction-diffusion equations for protein spatio-temporal signalling rising from protein-protein (and/or protein-mRNA) interactions in the two compartments of a cell, nucleus and cytoplasm, and from migration of the species in and between these compartments. Since it is impossible in general to derive analytical solutions for such coupled systems, numerically, passing to a weaker notion of solution, it can be proved to exist even in Lipschitz domains (as cells are assumed to be in our approach). With the Rothe method in hand, examples of such spatio-temporal oscillatory models, particularly, the Leloup-Goldbeter (β2GPI) antibodies, Anti-phospholipid antibodies (aPL) model and ATM-apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b-Wip1-Mdm2 model in individual cells are shown, numerically solved and illustrated. Let us mention here that works on reaction-diffusion models applied to intracellular biology had been rather scarce until recently. Among previous studies [74,89-92], it is worth pointing out the Ph D thesis of A. Serafini [93] that gives a different (theoretical and computational) approach to the reaction-diffusion equations applicable in modelling protein networks.The paper concludes with final remarks. In this work, we explore the MPNPE approach for modeling equilibrium and transport properties of ionic solutions in realistic three-dimensional geometries subject to realistic applied potentials on an in silico Reaction-diffusion Chemoproteomic-aided Molecular designed, polypharmacophoric ligand (CAMSPCPL) targeted on the conserved (β2GPI) antibodies, Anti-phospholipid antibodies (aPL) or from (PAPP-A; gi: 38045915) and APOH apolipoprotein H post-trancripts domains to Germline low energetics.

Materials and Methods

Reaction-diffusion systems on a 3D pattern regular expression methodology

Our methodology is inspired on the well-established PROSITE syntax commonly used in protein sequence analysis [94,95], which consists of a collection of pre-defined sequence profiles (defining protein families or functional motifs) that is used to find matches in protein sequence databases. In particular, our methodology involves two technical contributions: a 3D pattern syntax and a structural pattern search algorithm. The 3D pattern syntax developed here uses the prosite pattern syntax as starting point. It contains sequence patterns defining functionalities at residue level, and a variety of structural connectors defining distance and angle ranges and their inter-relationships. This syntax allows multiple sequence patterns to be joined using flexible structural connectors. The combination of all these syntax elements confers flexibility in the definition of sequence and structural constraints representing desired functionalities in each individual syntax query. A sequence pattern consists of a definition of one or a combination of amino acids. The later is represented with square brackets "[]" for allowed residues and curly brackets "{}" for forbidden residues. The geometric relationship between sequence patterns in three-dimensional space is established by using geometrical descriptors, which we call structural connectors.

A structural connector is defined in terms of a range of values:

i) Distances (in Angstroms) between the alpha carbons or a functional center of the side chain of amino acids, the latest denoted as pseudo atoms roughly equivalent to the center of mass of the side chain and inspired in previous work of Bahar [96];

ii) Angles (in degrees) formed by the corresponding peptide sequence backbones. The use of pseudo-points represents an advantage over calculating all pairwise distances, which would considerably increase the calculation time (about 16 times). Distances between alpha carbons or pseudo-points and angles between alpha carbons are specified in the search interface by selecting either alpha-carbons or pseudo-points for each search. A structural connector is denoted by a pair of angle brackets "< >". Structural constraints can be combined by introducing a pipe-line "|" between connectors. The connection between two sequence patterns and a structural connector is defined by a hyphen character (i.e. "-"). A cyclic pattern is built by adding this hyphen at the end of the last sequence pattern or structural connector.

A particular 3D pattern syntax query defining geometric relationships of desired functionalities is used as input by our structural pattern search algorithm. The structural pattern search algorithm operates across multiple stages designed to minimize the computational overhead of querying large protein structure datasets (i.e. the Protein Data Bank PDB) directly.

First, and as a pre-filtering step, the search engine utilizes a dataset of sequences (extracted from the 3D structures to be analyzed; i.e. PDB) for candidate hits matching. A candidate hit is any protein sequence that has each of the specified sequence elements or sub-sequences in the pattern. Sub-sequences should not overlap.

A data-structure of all available sub-sequences in a given protein (often there are multiple matches for each sub-sequence) is built by considering all hits obtained in the previous step. This data-structure is then used in the next stage.

The engine verifies that there is at least one set of sub-sequences that matches the required structural constraints specified in the 3D pattern syntax query. As the combinatorial space can be extremely large, the engine makes use of a recursive backtracking algorithm to minimize redundant computation and permit early termination when one of the required sub-sequences has been found to fail all structural constraints.

As multiples matches could be found for the same structure, and in order to improve computational efficiency and to avoid large numbers of very minor variations, we add the additional constraint that multiple matches in the same structure cannot overlap any components. In addition, we run the structural pattern search algorithm over a redundancy reduced data set. When matches are found, we are able to manually explore similar structures for equivalent matches.

All algorithms developed to interrogate the protein structure space in a systematic manner searching for patterns in databases have to make a trade-off between three key areas:

1) The pattern expressivity (i.e. range of functional motifs that can be captured)

2) The ease of pattern definition

3) The computational efficiency

In our approach, we scarify some of the first feature in order to achieve better performance on the latter two. Our methodology presents a number of key differences with other comparable computational techniques such as RAMOT-3D [96], ASSAM [97] and IMAAAGINE [98] and overcomes some of their limitations. Typically, in these tools the search syntax does not allow for variations in the identities of residues in particular positions as our does. In our case, the patterns are inspired on the PROSITE syntax and therefore defined using amino acid sequence elements. As result easy and flexible patterns can be designed instead of using graphs. Furthermore, our patterns can have arbitrary numbers of amino acids containing distance and angle tolerances described for each connection instead of being globally defined. Besides, the chemical properties in our patterns are defined in our case only through sets of amino acids and not via many pre-defined sets.

Computational structure-based design and simulation

We used the 3D crystal structure of human apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b for our calculations (PDB This structure contains a protein dimer and two receptor molecules. Due to the two-fold axis symmetry in this complex, only one of the two apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b domains and a receptor molecule were considered for our studies. Analysis of relevant binding residues in the apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b/a. To determine the receptor residues being the major contributors to the protein binding energy, molecular dynamics calculations were carried out with the apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b/apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_bR1 complex. AMBER11 and AMBER12 packages [97] were used for the simulations, and the MM-PBSA method [98,99] as implemented in AMBER11 was used to obtain per residue energy contributions. The ff99SB force field [100] and standard protocols as implemented in the AMBER11 and AMBER12 packages were used (see below) [97]. The apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b/apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_bR1 complex was solvated in a truncated octahedral box of TIP3P water molecules and neutralized with Cl- counterions.

Molecular dynamics simulations were preceded by two energy-minimization steps:

i) Only the solvent and ions were relaxed;

ii) The entire system was minimized.

The system was then heated up from 0 K to 300 K in 50 ps. Weak position restraints (10 kcal/mol•Å2) were used for the whole system in the canonical ensemble (NVT). Langevin temperature coupling with a collision frequency $\gamma = 1$ ps-1 was used at this step. The system was equilibrated without restraints at 300 K during 100 ps in the isothermal-isobaric ensemble (NPT) followed of other 100 ps in the canonical ensemble (NVT). The Langevin thermostate and Berendsen barostat under periodic boundary conditions were used during the equilibration steps. After this, a total of 20 ns MD production without restraints was carried out at 300 K in the canonical ensemble (NVT) using the Langevin thermostat. The SHAKE algorithm was used to constrain all bonds involving hydrogen atoms. A time step of 2 fs was used. A cutoff of 8 Å was applied to treat the non-bonded interactions, and the Particle Mesh Ewald (PME) method was used to treat long-range electrostatic interactions. MD trajectories were recorded every 2 ps. Trajectories were visualized with VMD [101] and evaluated in terms of intermolecular H-bonds using the PTRAJ and CPPTRAJ modules implemented in AMBER. The criterion used to consider a dynamic hydrogen bond formation was to be found at least 10% of the simulation time using a distance acceptor-donor cutoff of 3.5 Å and a 120° angle cut off.

Mimetic refinement and analysis

In each step of the lactam bridge design strategy, a brief minimization was performed with MOE (Amber12: EHT force field

and default parameters were used) [102]. Molecular dynamics simulations of the complex of apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b with each of the designed apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_bR1 mimetics were performed using the ff99SB force field [97] and standard protocols as implemented in the AMBER11 and AMBER12 packages (see below) [103]. The mimetic residues Lys and Asp involved in lactam bridges were parametrized according to the ff99SB force field, and RESP charges were derived at the HF/6-31G* calculation level using Gausian09 [104-106]. Same strategy was followed for parameters of the chemical groups introduced in the optimization process that were not included in the ff99SB force field. Each protein-mimetic complex was solvated in a truncated octahedral box of TIP3P water molecules and neutralized with Cl- counterions.

Molecular dynamics simulations were preceded by two energy-minimization steps:

i) Only the solvent and ions were relaxed

ii) The entire system was minimized with low position restraints (10 kcal/mol•Å2) for the helical backbone section of the mimetic. The system was then heated up from 0 K to 300 K in 20 ps. Weak position restraints (10 kcal/mol•Å2) were used for the protein and the mimetic in the canonical ensemble (NVT). Langevin temperature coupling with a collision frequency $\gamma = 1$ ps-1 was used at this step. Along three equilibration steps of 500 ps each, the helical backbone position restraints and also two distance restraints between protein and mimetic (apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_bQ38/MR7 and apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_bK138/MY4) were consecutively decreased (10, 5, and 2 kcal/mol•Å2, respectively). This was carried out in the isothermal-isobaric ensemble (NPT) using the Langevin thermostate and Berendsen barostat under periodic boundary conditions. The system was further equilibrated during 1ns without restraints at 300 K under same conditions. After this, a 10 ns MD production was carried out at 300 K in the canonical ensemble (NVT) using the Langevin thermostat. The SHAKE algorithm was used to constrain all bonds involving hydrogen atoms. A time step of 2 fs was used. A cutoff of 10 Å was applied to treat the nonbonded interactions, and the Particle Mesh Ewald (PME) method was used to treat long-range electrostatic interactions. MD trajectories were recorded every 2 ps. Trajectories were visualized with VMD [368] and evaluated in terms of intermolecular H-bonds (including bridging water molecules at the interface between the designed mimetics and apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b) and RMSD using PTRAJ and CPPTRAJ modules implemented in AMBER. The criterion used to consider a dynamic hydrogen bond formation was to be found at least 10% of the total simulation time. Water molecules connecting functionalities (bridging waters) were also considered for hydrogen bond analysis using a distance acceptor-donor cutoff of 3.5 Å and angle cutoff of 1200. A structural bridging water was considered to be participating in H-bond in protein-mimetic complex when was found at least 20% of the total simulation time. Energy decomposition per residue and binding free energy post-processing analysis of the trajectories Table 2 were performed in implicit solvent using the MM-PBSA method [98,99] as implemented in AMBER11. The same method was applied to study the effect of alanine mutation on mimetic residues R3, Y4 and R7 of M1, M4 and M6 in complex with apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b [98,99,107]. Data analysis was carried out with the R-package [108]. In order to ensure that main conclusions inferred from the 10 ns MD simulations maintained valid our strategy design, a further 20ns extension was carried out for the optimized molecules (i.e. apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b/M2-M6 complexes).

GRID calculations

GRID [109,110] version 22 was used to predict energetically favorable interactions between apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b (PDB ID 1J7V) and the following chemical probes: N2 = (sp2 amine NH2 cation), N3+ (sp3amine NH3 cation), COO- (aliphatic carboxylate), and between the unbound apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b (PDB ID 2ILK) and the chemical probes: OH2 (water) and C3 (methyl CH3 group). The GRID box dimensions were set to 51 Å x 58 Å x 46 Å in order to cover the whole protein for the calculations. A grid spacing of 1 Å was applied. The rest of GRID input parameters values were used as default. Structure-based modeling of mapolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b C149Y. The crystal structure of human apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b at high resolution (PDB ID 2ILK (1.6 Å)) [111] and its complex structure with apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_bR1 (PDB ID 1J7V (2.9 Å) [112] were taken as structural templates. Human and mouse proteins and their R1 receptors share high sequence similarities (88% and 73%, respectively). The mouse receptor was modeled following the same procedure. Spatio-temporal dynamics of a variety of proteins is, among other things, regulated by post-translational modifications of these proteins [113-115]. Such modifications can thus influence stability and biochemical activities of the proteins, activity and stability of their upstream targets within specific signalling pathways. Commonly used mathematical tools for such protein-protein (and/or protein-mRNA) interactions in single cells, namely, Michaelis-Menten and Hill kinetics, yielding a system of ordinary differential equations, are extended here into (non-linear) partial differential equations by taking into account a more realistic spatial representation of the environment where these reactions occur [116-121]. In the modelling framework under consideration, all interactions occur in a cell divided into two compartments, the nucleus and the cytoplasm, connected by the semipermeable nuclear membrane and bounded by the impermeable cell membrane [122-125]. Passive transport mechanism, modelled by the so-called Kedem-Katchalsky boundary conditions, is used here to represent migration of species throughout the nuclear membrane [126-130].

Nonlinear systems of partial differential equations are solved by the semi-implicit Rothe method. Examples of two spatial oscillators are shown [131-133]. Namely, these are the circadian rhythm for concentration of the (β2GPI) antibodies, Anti-phospholipid antibodies (aPL) protein in Infertility and oscillatory dynamics observed in the activation and regulation of the apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b protein following DNA damage in mammalian cells [134-137]. Cellular responses are controlled either by one particular and functionally active protein or by several proteins whose activation and activity towards other proteins in highly specific situations depend on other factors and conditions which cells are exposed to such processes may be also different from cell to cell; they can be influenced by extracellular factors such as light, heat, abundance of stress agents and duration of their action, growth conditions, available resources and many others [138-140]. On a molecular basis, roles of proteins in specific networks are influenced by protein-protein interactions either through post-translational modifications (by attaching a phosphate or acetyl group, ubiquitin to a protein, etc.), or by various compound formation [141]. An interesting example is the signalling network of the protein apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b that may elicit life and death decisions in a cell [142]. Activity of apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b in such situations (meaning not only transcriptional activity towards pro-arrest, pro-apoptotic and pro-survival target genes) is closely controlled by post-translational modifications (such as phosphorylation or ubiquitination) and interactions with other proteins [143-145].

Kinetics of enzyme reactions

Let us briefly summarise basic mathematical ideas used to model an enzyme reaction

$$S+En\rightleftharpoons k-1k1SEn\rightarrow k2En+P$$

where a substrate S reacts with an enzyme En forming a complex SEn with a rate k1. In this complex, the enzyme En converts S into a product protein P with a kinetic rate k2 [146]. The enzyme En is released and available for further reactions. The first reaction between S and En is reversible, meaning that the complex SEn can eventually fall apart with a reverse rate of reaction k- 1, whilst the second reaction is assumed to proceed in one direction only [147].

The law of mass action for the concentrations of species denoted subsequently by s = [S],e = [En],c = [SEn] and p = [P] gives four equations, one for each species. These are

$$dsdt =k-1c-k1se,dedt=k-1+k2c-k1se,\ dcdt=k1se-k-1+k2c,\ dpdt=k2c$$

and we can assume initial conditions s(0) = s0, e(0) = e0, c(0) = 0 and p(0) = 0.

Note that the last equation can be easily solved whenever c is determined. In this case $p(t) = k2\int 0tc(\tau)d\tau$ and so the last equation can be omitted from further considerations. Note also that dedt+dcdt=0 and so e + c = e0 where e0 is the total amount of available enzyme. By substituting e = e0 - c into the equations for s and c, we can write

$$dsdt =k-1+k1sc-k1se0-(1)$$

$$dcdt =k1se0-k1s+k-1+k2c-(2)$$

The quasi-steady-state approximation (QSSA) can be finally applied to eliminate the equation for the complex. In particular, one can approximate the rate of change of c by zero, i.e. dcdt≈0. Although equality is not always true during the reaction, see for example [146], by taking the right hand side of Eq. (2) to be zero, one can write c in terms of s and such term substitution into Eq. (1) finally yields

$$dsdt=-k2e0sK2+swithK2=k-1+k2k1-(3)$$

where k2 is the kinetic rate of the reaction and K2 is affinity constant.

A basic assumption for the QSSA is that the concentration of enzyme is smaller than the concentration of substrate (e.g., for ε = e0/s0 small, typically, between 10-2 and 10-7, [147], the assumption that is not usually true in our models since enzymes under consideration may be present in large quantities. The QSSA can however be found valid also in the case when the concentration of the intermediate complex SEn is immediately consumed, i.e. it is always small compared with the total concentration of the substrate [146,147]. For a rigorous mathematical treatment of the QSSA from the point of view of perturbation theory, see [146]. Eq. (3) is thought to represent the total loss of the substrate S in the reaction, and thus, by using conservation of mass, the same term with opposite sign can be written as the total gain of the product P. In practical simulations, e0 in Eq. (3) is often replaced by the actual concentration e.

Example 2.1

In a model discussed in Section 3.2, we will consider the protein apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b, its activation by the kinase ATMp (phosphorylation of apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b by ATM giving apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b p) and regulated by the phosphatase Wip1 and the E3 ligase Mdm2 (dephosphorylation of apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b p by Wip1 and ubiquitination of apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b by Mdm2, respectively). All these modifications can be represented as enzyme reactions, i.e.

Apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform

CRA_b+ATMp⇌k-atmkatmComplex→kph1ATMp+apolipoprotein

H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b pphosphorylationof apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b byATMp,apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b p+Wip1⇌k-wip1kwip1Complex→kdph1Wip1+apolipoprote in H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b dephosphorylationof apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b p by

Wip1, p53+Mdm2⇌k-mdm2kmdm2Complex→kubMdm2+p53ububiquitinationofp53byMdm2

Following the approach described above, one finally arrives at the equations for p53 and its phosphorylated version p53p, in particular,

dp53dt=-kph1ATMpp53Kph1+p5 ⁻⁻ 3phosphorylationofp53byATMp+kdph1Wip1p53pKdph1+p53p⁻⁻de

phosphorylationofp53pbyWip1-kubMdm2p53Kub+p53⁻⁻ubiquitinationofp53byMdm2,dp53pdt=kph1ATMpp53Kph1+p53⁻⁻ phosphorylationofp53byATMp-kdph1Wip1p53pKdph1+p53p⁻⁻dephosphorylationofp53pbyWip1 with specific kinetic maximum velocity rates kx and affinity constants Kx. The situation is slightly complicated whenever monomerisation and polymerisations of a protein are modelled as enzyme reactions, i.e.

SN+En⇌k-ikiComplex→kaEn+NPNS+En⇌k-jkjComplex→kbEn+PN

Kinetics of enzyme reactions using the law of mass action and the QSSA is sometimes collectively called Michaelis-Menten kinetics [146,147].

In the models presented below, we will also consider the expression of genes, particularly transcription of (β2GPI) antibodies, Anti-phospholipid anti-bodies (aPL), mdm2 and wip1 into messenger RNA and its translation into proteins. Whilst translation is usually modelled as linear gain of the concentration of translated proteins, transcription can be modelled as enzyme reaction; however, transcription factors often express cooperative binding to the active sites of DNA (considered to be the enzyme in the reaction), i.e. binding of one substrate molecule affects binding of subsequent substrate molecules, as it is in the case of p53 that bind DNA as tetramers and thus four molecules of p53 cooperate whilst performing the transcription job [148,149]. In such kinetics, sometimes called non-Michaelis-Menten kinetics or Hill kinetics, the Hill equation can be used to describe the degree of cooperativity quantitatively [147,150]. In particular, transcription of mRNA can be described by a Hill function

dmRNAdt = ktTFnKtn+TFn

With Hill coefficient n, rate of transcription kt, affinity constant Kt and a transcription factor TF. According to [148], see also [68] chap. 1, the Hill coefficient is equal to the number of binding sites, which, for example, in the case of tetrameric p53 binding DNA, is n = 4.

Finally, natural degradation terms will be modelled here by linear loss terms or by Hill functions of coefficient 1. In the p53 model there is also a decay term in the nuclear and cytoplasmic p53 equation coming from ubiquitination of p53 by Mdm2, thus forming Michaelis-Menten kinetics discussed above. This is because p53 is degraded by the ubiquitin-dependent degradation machinery in many conditions [85]; for the sake of simplicity, we will not introduce any special equation for ubiquitinated p53 since it would consist only of degradation terms for this species.

Reaction-diffusion models for protein cellular signalling

In PDEs, the protein dynamics in cells is particularly endowed with diffusivity rates and with some transmission conditions on the nuclear and cell (possibly other) membranes to represent, respectively, migration of species in and their exchange between the cellular compartments and between cells themselves. In the two sections that follow we show two examples of oscillators given by reaction-diffusion PDEs [118,119].

In cellular settings, a general coupled reaction diffusion system may be written as follows,

dudt-divD∇u=f(u)in(0,T)×Ω1,dvdt-divD∇v=g(v)in(0,T)×Ω2,+boundaryconditionsonΓ1andΓ2-(4)

where u = u(t,x) : [0,T] × Ω1]N → RN and v = v(t,x) : [0,T] × Ω2]N → RN are the nuclear and cytoplasmic concentrations of N chemicals, T > 0, N > 0, f:RN → RN and g:RN → RN collect protein-protein and protein-mRNA reaction terms, and production/degradation terms for all N species as they are discussed in Section 2, D is a diagonal N-by-N matrix with the diffusivities on the diagonal, and div is the divergence operator. The domains Ω1 and Ω2 represent the nucleus and the cytoplasm, respectively, assumed to be Lipschitz domains [46,151], Γ1 is the nuclear membrane and Γ2 is the cell membrane. The geometry of a cell under consideration that in our very simplified framework represents a cell as big as 10-100μm or so, which is able to proliferate and which can have heterogeneous shape. However, we exclude cells of complicated structures and morphologies such as neurons or muscle cells. Cell scheme: the nucleus Ω1, the cytoplasm Ω2, the nuclear membrane Γ1 and the cell membrane Γ2; n1 and n2 are the unit normal vectors oriented outward from Ω1 and Ω2, respectively.

As previously outlined, we will consider zero-flux boundary conditions on the outer cell membrane Γ2, i.e. ∂vi∂n2=0 for i = 0,1..., N-1 where n2 is the unit normal vector oriented outward from the cell. Although species with molecular weight over 40 kDa usually use active transport mechanisms to be translocated from one compartment to another, for the sake of simplicity, only passive transport process driven by the difference in concentrations at both sides of the nuclear membrane, modelled by the so-called Kedem-Katchalsky BCs, is used here [74,113,123,124]. In particular, export/import of the nuclear species is modelled by

Di∂ui∂n1=-pivi-ui-(5)

and export/import of the cytoplasmic species by

Di∂vi∂n1=-piui-vi-(6)

where n1 is the unit normal vector oriented outward from the nucleus (therefore, the BCs in Eq. (6) come with a different sign), and pi are the permeabilities of the membrane for each species i = 0,1..., N - 1. Recall that if a chemical is assumed to migrate between the compartments in one direction only, e.g. mRNA is supposed to move from the transcription sites in the nucleus to the cytoplasm and not the other way round, then this can easily be modelled by omitting either vi or ui in Eqs. (5) and (6), cf. (9)-(11) and (14) below.

In addition to the boundary conditions, we will assume initial conditions for each species i = 0,1..., N - 1 to be non-negative functions ui(t = 0,x) = ui0(x) ≥ 0 and vi(t = 0, x) = vi0(x) ≥ 0 belonging to L2(Ω1) and L2(Ω2), respectively.

The Leloup-Goldbeter Model

$$\frac{d(\mathrm{mRNA})}{dt} = \frac{a_1}{A_1 + k_1 (P)} - \delta_1,$$

$$\frac{d(P)}{dt} = a_2(\mathrm{mRNA}) - \delta_2,$$

$$S + \mathrm{En} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \mathrm{SEn} \xrightarrow{k_2} \mathrm{En} + P$$

$$\frac{ds}{dt} = k_{-1}c - k_1\,\mathrm{se},\ \frac{de}{dt} = (k_{-1} + k_2)c - k_1\,\mathrm{se},$$

$$\frac{dc}{dt} = k_1\,\mathrm{se} - (k_{-1} + k_2)c,\ \frac{dp}{dt} = k_2 c,$$

$$\frac{de}{\mathrm{dt}} + \frac{dc}{dt} = 0$$

$$\frac{ds}{dt} = (k_{-1} + k_1 s)c - k_1 s e_0,$$

$$\frac{dc}{dt} = k_1 s e_0 - (k_1 s + k_{-1} + k_2)c.$$

$$\frac{dc}{dt} \approx 0$$

$$\frac{ds}{dt} = -k_2 e_0 \frac{s}{K_2 + s} \mathrm{with} K_2 = \frac{k_{-1} + k_2}{k_1}$$

$$p53 + \mathrm{AT}M_p \underset{k_{-\mathrm{atm}}}{\overset{k_{\mathrm{atm}}}{\rightleftharpoons}} \mathrm{Complex} \xrightarrow{k_{\mathrm{ph1}}} \mathrm{AT}M_p + p53_p\,(\mathrm{phosphorylation\,of}\,p53\,\mathrm{by\,AT}M_p)$$

$$p53_p + \mathrm{Wip1} \underset{k_{-\mathrm{wip1}}}{\overset{k_{\mathrm{wip1}}}{\rightleftharpoons}} \mathrm{Complex} \xrightarrow{k_{\mathrm{dph1}}} \mathrm{Wip1} + p53(\mathrm{dephosphorylation\,of}\,p53_p\,\mathrm{by\,Wip1}$$

$$p53 + \mathrm{Mdm2} \underset{k_{-\mathrm{mdm2}}}{\overset{k_{\mathrm{mdm2}}}{\rightleftharpoons}} \mathrm{Complex} \xrightarrow{k_{\mathrm{ub}}} \mathrm{Mdm2} + p53_{\mathrm{ub}}(\mathrm{ubiquitination\,of}\,p53\,\mathrm{by\,Mdm2}).$$

$$\frac{d(p53)}{dt} = \underbrace{-k_{\mathrm{ph1}}(\mathrm{AT}M_p)\frac{(p53)}{K_{\mathrm{ph1}} + (p53)}}_{\mathrm{phosphorylation\,of}\,p53\,\mathrm{by\,AT}\,Mp} + \underbrace{k_{\mathrm{dph1}}(\mathrm{Wip1})\frac{(p53_p)}{K_{\mathrm{dph1}} + (p53_p)}}_{\mathrm{dephosphorylation\,of}\,p53_p\,\mathrm{by\,Wip1}} - \underbrace{k_{\mathrm{ub}}(\mathrm{Mdm2})\frac{(p53)}{K_{\mathrm{ub}} + (p53)}}_{\mathrm{ubiquitination\,of}\,p53\,\mathrm{by\,Mdm2}},$$

$$\frac{d(p53_p)}{dt} = \underbrace{k_{\mathrm{ph1}}(\mathrm{AT}M_p)\frac{(p53)}{K_{\mathrm{ph1}} + (p53)}}_{\mathrm{phosphorylation\,of}\,p53\,\mathrm{by\,AT}\,Mp} - \underbrace{k_{\mathrm{dph1}}(\mathrm{Wip1})\frac{(p53_p)}{K_{\mathrm{dph1}} + (p53_p)}}_{\mathrm{dephosphorylation\,of}\,p53_p\,\mathrm{by\,Wip1}}$$

$$S(N) + \mathrm{En} \underset{k_{-i}}{\overset{k_i}{\rightleftharpoons}} \mathrm{Complex} \xrightarrow{k_a} \mathrm{En} + \mathrm{NP}$$

$$\mathrm{NS} + \mathrm{En} \underset{k_{-j}}{\overset{\mathrm{kj}}{\rightleftharpoons}} \mathrm{Complex} \xrightarrow{k_b} \mathrm{En} + P(N)$$

$$\frac{d(\mathrm{mRNA})}{dt} = k_t \frac{(\mathrm{TF})^n}{K_t^n + (\mathrm{TF})^n}$$

$$\frac{du}{dt} - \mathrm{div}(D\nabla u) = f(u)\mathrm{in}(0, T) \times \Omega_1,$$

$$\frac{dv}{dt} - \mathrm{div}(D\nabla v) = g(v)\mathrm{in}(0, T) \times \Omega_2,$$

$$+\mathrm{boundary\,conditions\,on}\,\Gamma_1\,\mathrm{and}\,\Gamma_2$$

$$\frac{\partial v_i}{\partial n_2} = 0$$

$$-D_i \frac{\partial u_i}{\partial n_1} = -p_i(v_i - u_i)$$

$$\frac{d(\mathrm{FRQ})^{(c)}}{dt} = k_s(\mathrm{FRQmRNA}) - V_d \frac{(\mathrm{FRQ})^{(c)}}{K_d + (\mathrm{FRQ})^{(c)}} - k_1[\mathrm{FRQ}]^{(c)} + k_2[\mathrm{FRQ}]^{(n)},$$

$$\frac{d(\mathrm{FRQ})^{(n)}}{dt} = k_1(\mathrm{FRQ})^{(c)} - k_2[\mathrm{FRQ}]^{(n)}.$$

$$u_0 = (\mathrm{FRQmRNA})^{(n)}, v_0 = (\mathrm{FRQmRNA})^{(c)}, u_1 = (\mathrm{FRQ})^{(n)} \text{and} v_1 = (\mathrm{FRQ})^{(c)}$$

$$\frac{\partial u_0}{\partial t} - D_0 \Delta u_0 = V_s \frac{K^r}{K^r + u_1^r} - V_m^n \frac{u_0}{K_m^n + u_0}, \quad \frac{\partial u_1}{\partial t} - D_1 \Delta u_1 = -V_d^n \frac{u_1}{K_d^n + u_1},$$

$$\frac{\partial v_0}{\partial t} - D_0 \Delta v_0 = -V_m^c \frac{v_0}{K_m^c + v_0}, \quad \frac{\partial v_1}{\partial t} - D_1 \Delta v_1 = k_s v_0 \chi_{\mathrm{tra}} - V_d^c \frac{v_1}{K_d^c + v_1},$$

$$-D_1 \frac{\partial u_1}{\partial n_1} = -p_1(v_1 - u_1) = -D_1 \frac{\partial v_1}{\partial n_1}.$$

$$-D_1 \frac{\partial u_1}{\partial n_1} = -p_1 v_1 = -D_1 \frac{\partial v_1}{\partial n_1}.$$

$$-D_0 \frac{\partial u_0}{\partial n_1} = p_0 u_0 = -D_0 \frac{\partial v_0}{\partial n_1}$$

$$\frac{\partial v_i}{\partial n_2} = 0$$

$$u_0 = (p53)^{(n)}, u_1 = [\mathrm{Mdm2}]^{(n)}, u_2 = [\mathrm{Mdm2mRNA}]^{(n)}, u_3 = [p53_p]^{(n)},$$

$$u_4 = (\mathrm{AT}M_p)^{(n)}, u_5 = [\mathrm{Wip1}]^{(n)}, u_6 = [\mathrm{Wip1mRNA}]^{(n)},$$

$$v_0 = (p53)^{(c)}, v_1 = [\mathrm{Mdm2}]^{(c)}, v_2 = [\mathrm{Mdm2mRNA}]^{(c)}, v_3 = [p53_p]^{(c)},$$

$$v_4 = (\mathrm{AT}M_p)^{(c)}, v_5 = [\mathrm{Wip1}]^{(c)}, v_6 = [\mathrm{Wip1mRNA}]^{(c)}.$$

$$\frac{du_0}{dt} = k_{\mathrm{dph1}} u_5 \frac{u_3}{K_{\mathrm{dph1}} + u_3} - k_{\mathrm{ub}} u_1 \frac{u_0}{K_{\mathrm{ub}} + u_0} - k_{\mathrm{ph1}} u_4 \frac{u_0}{K_{\mathrm{ph1}} + u_0} - p_0 V_r(u_0 - v_0)$$

$$\frac{dv_0}{dt} = k_S - k_{\mathrm{ub}} v_1 \frac{v_0}{K_{\mathrm{ub}} + v_0} - p_0(v_0 - u_0) - \delta_0 v_0.$$

$$\frac{du_1}{dt} = -p_1 V_r(u_1 - v_1) - \delta_1 u_1 \qquad \frac{dv_1}{dt} = k_{\mathrm{tm}} v_2 - p_1(v_1 - u_1) - \delta_1 v_1$$

$$\frac{du_2}{dt} = k_{S\mathrm{pm}} \frac{u_3^4}{K_{S\mathrm{pm}}^4 + u_3^4} - p_2 V_r u_2 - \delta_2 u_2 \qquad \frac{dv_2}{dt} = p_2 u_2 - k_{\mathrm{tm}} v_2 - \delta_2 v_2$$

$$\frac{du_3}{dt} = k_{\mathrm{ph1}} u_4 \frac{u_0}{K_{\mathrm{ph1}} + u_0} - k_{\mathrm{dph1}} u_5 \frac{u_3}{K_{\mathrm{dph1}} + u_3} \qquad \frac{dv_3}{dt} = 0$$

$$\frac{du_4}{dt} = k_{\mathrm{ph2}} E \frac{\mathrm{AT}M_{\mathrm{TOT}} - u_4}{K_{\mathrm{ph2}} + 1/2(\mathrm{AT}M_{\mathrm{TOT}} - u_4)} - 2k_{\mathrm{dph2}} u_5 \frac{u_4^2}{K_{\mathrm{dph2}} + u_4^2} \qquad \frac{dv_4}{dt} = 0$$

$$\frac{du_5}{dt} = p_5 V_r v_5 - \delta_5 u_5 \qquad \frac{dv_5}{dt} = k_{\mathrm{tw}} v_6 - p_5 v_5 - \delta_5 v_5,$$

$$\frac{du_6}{dt} = k_{S\mathrm{pw}} \frac{u_3^4}{K_{S\mathrm{pw}}^4 + u_3^4} - p_6 V_r u_6 - \delta_6 u_6 \qquad \frac{dv_6}{dt} = p_6 u_6 - k_{\mathrm{tw}} v_6 - \delta_6 v_6$$

$$\frac{\partial u_0}{\partial t} - D_0 \Delta u_0 = k_{\mathrm{dph1}} u_5 \frac{u_3}{K_{\mathrm{dph1}} + u_3} - k_{\mathrm{ub}} u_1 \frac{u_0}{K_{\mathrm{ub}} + u_0} - k_{\mathrm{ph1}} u_4 \frac{u_0}{K_{\mathrm{ph1}} + u_0},$$

$$\frac{\partial u_1}{\partial t} - D_1 \Delta u_1 = -\delta_1 u_1,$$

$$\frac{\partial u_2}{\partial t} - D_2 \Delta u_2 = k_{S\mathrm{pm}} \frac{u_3^4}{K_{S\mathrm{pm}}^4 + u_3^4} - \delta_2 u_2,$$

$$\frac{\partial u_3}{\partial t} - D_3 \Delta u_3 = k_{\mathrm{ph1}}\, u_4 \frac{u_0}{K_{\mathrm{ph1}} + u_0} - k_{\mathrm{dph1}}\, u_5 \frac{u_3}{K_{\mathrm{dph1}} + u_3},$$

$$\frac{\partial u_4}{\partial t} - D_4 \Delta u_4 = k_{\mathrm{ph2}} E \frac{\mathrm{AT}\, M_{\mathrm{TOT}} - u_4}{K_{\mathrm{ph2}} + \frac{1}{2}(\mathrm{AT}\, M_{\mathrm{TOT}} - u_4)} - 2k_{\mathrm{dph2}}\, u_5 \frac{u_4^2}{K_{\mathrm{dph2}} + u_4^2},$$

$$\frac{\partial u_5}{\partial t} - D_5 \Delta u_5 = -\delta_5 u_5,$$

$$\frac{\partial u_6}{\partial t} - D_6 \Delta u_6 = k_{\mathrm{Spw}} \frac{u_3^4}{K_{\mathrm{Spw}}^4 + u_3^4} - \delta_6 u_6,$$

$$\frac{\partial v_0}{\partial t} - D_0 \Delta v_0 = k_S \chi_{\mathrm{bp}} - k_{\mathrm{ub}}\, v_1 \frac{v_0}{K_{\mathrm{ub}} + v_0} - \delta_0 v_0,$$

$$\frac{\partial v_1}{\partial t} - D_1 \Delta v_1 = k_{\mathrm{tm}}\, v_2 \chi_{\mathrm{tra}} - \delta_1 v_1,$$

$$\frac{\partial v_2}{\partial t} - D_2 \Delta v_2 = -k_{\mathrm{tm}}\, v_2 \chi_{\mathrm{tra}} - \delta_2 v_2,$$

$$\frac{\partial v_3}{\partial t} - D_3 \Delta v_3 = 0,$$

$$\frac{\partial v_4}{\partial t} - D_4 \Delta v_4 = 0,$$

$$\frac{\partial v_5}{\partial t} - D_5 \Delta v_5 = k_{\mathrm{tw}}\, v_6 \chi_{\mathrm{tra}} - \delta_5 v_5,$$

$$\frac{\partial v_6}{\partial t} - D_6 \Delta v_6 = -k_{\mathrm{tw}}\, v_6 \chi_{\mathrm{tra}} - \delta_6 v_6$$

$$\frac{\partial v_0}{\partial t} - D_0 \Delta v_0 = k_S \chi_{\mathrm{bp}} - k_{\mathrm{ub}}\, v_1 \frac{v_0}{K_{\mathrm{ub}} + v_0} - \delta_0 v_0,$$

$$\frac{\partial v_1}{\partial t} - D_1 \Delta v_1 = k_{\mathrm{tm}}\, v_2 \chi_{\mathrm{tra}} - \delta_1 v_1,$$

$$\frac{\partial v_2}{\partial t} - D_2 \Delta v_2 = -k_{\mathrm{tm}}\, v_2 \chi_{\mathrm{tra}} - \delta_2 v_2,$$

$$\frac{\partial v_3}{\partial t} - D_3 \Delta v_3 = 0,$$

$$\frac{\partial v_4}{\partial t} - D_4 \Delta v_4 = 0,$$

$$\frac{\partial v_5}{\partial t} - D_5 \Delta v_5 = k_{\mathrm{tw}}\, v_6 \chi_{\mathrm{tra}} - \delta_5 v_5,$$

$$\frac{\partial v_6}{\partial t} - D_6 \Delta v_6 = -k_{\mathrm{tw}}\, v_6 \chi_{\mathrm{tra}} - \delta_6 v_6 .$$

$$\frac{du}{dt} - \mathrm{div}(D \nabla u) = f(u) \text{ and } \frac{dv}{dt} - \mathrm{div}(D \nabla v) = g(v)$$

$$u_i \in W_1^{1,2,2}(I, V_1, V_1^*) = \left(u \in L^2(I, V_1); \frac{du}{dt} \in L^2(I, V_1^*)\right)$$

$$\frac{du}{dt}$$

$$\frac{du}{dt} - A(u(t)) = \tilde{f}(t) \text{ and } \frac{dv}{dt} - B(v(t)) = \tilde{g}(t) \text{ for } a.a.\, t \in I$$

$$\text{with } u(0) = u_0,\ v(0) = v_0 \text{ and BCs on } \Gamma_1 \text{ and } \Gamma_2,$$

$$\widetilde{f}$$

$$\widetilde{g}$$

$$\widetilde{f}$$

$$\widetilde{g}$$

$$\widetilde{f}$$

$$\widetilde{g}$$

$$\overline{A}(z, \cdot): V_1 \to V_1^*$$

$$\overline{B}(z, \cdot): V_2 \to V_2^*$$

$$\frac{u_\tau^k - u_\tau^{k-1}}{\tau} + \overline{A}\left(u_\tau^k, u_\tau^{k-1}\right) = \widetilde{f}(k\tau) \text{and} \frac{v_\tau^k - v_\tau^{k-1}}{\tau} + \overline{B}\left(v_\tau^k, v_\tau^{k-1}\right) = \widetilde{g}(k\tau), k = 1, \ldots, T/\tau,$$

$$\overline{B}(v, v) = B(v)$$

$$-D\frac{\partial u_\tau^k}{\partial n_1} = -p\left(v_\tau^k - u_\tau^k\right) = -D\frac{\partial v_\tau^k}{\partial n_1} \text{on} \Gamma_1 \text{and} -D\frac{\partial v_\tau^k}{\partial n_2} = 0 \text{on} \Gamma_2.$$

$$-D_0\frac{\partial u_{0,\tau}^k}{\partial n_1} = p_0 u_{0,\tau}^k = -D_0\frac{\partial v_{0,\tau}^k}{\partial n_1},$$

$$-D_1\frac{\partial u_{1,\tau}^k}{\partial n_1} = -p_1 v_{1,\tau}^k = -D_1\frac{\partial v_{1,\tau}^k}{\partial n_1}.$$

Reaction-diffusion systems for spatio-temporal intracellular protein networks-A beginner's guide with two examples

The semi-implicit rothe method for numerical applications

The semi-implicit Rothe method [138] can be used when one has to solve a coupled reaction-diffusion model (4) in cells as the ones previously discussed in Sections 3.1 and 3.2 as well as it can be used to prove theoretically existence of such solutions. Let T > 0 and I = [0,T] be a fixed time interval. Without going into details (a special article [122] is dedicated to more precise description of the method and rigorous proofs of statements claimed here), let us mention that we are looking for solutions u = [u0,u1,...,uN - 1]T and v = [v0,v1,...,vN - 1]T to the coupled reaction-diffusion system

dudt-divD∇u=fuanddvdt-divD∇v=gv-(15)

defined in two time-space cylinders I × Ω1 and I × Ω2 connected on a common boundary I × Γ1 by Robin-like boundary conditions (Kedem-Katchalsky BCs) and satisfying zero-flux BCs on I × Γ2 and some initial conditions at time t = 0. Recall that in d-dimensional case, d ≥ 1, Ω1 and Ω2 are open and bounded sets in Rd with Lipschitz boundaries Γ1 = ∂Ω1∩∂Ω2 and Γ2 = ∂Ω2 - Γ1, see also Figure for a schematic representation of the domains. Solutions ui for i = 0,1,... N-1 is supposed to belong to a standard.

Sobolev-Bochner space

ui∈W11,2,2IV1V1∗ =u∈L2IV1;dudt∈L2IV1∗

With V1 = W1,2(Ω1) and V1∗ dual to V1. Similarly, vi∈W21,2,2(I;V2,V2∗) with V2 = W1,2(Ω2) and V2∗ dual to V2 for all i = 0,1,... N-1. By dudt we understand the derivative of u in sense of distributions. We will refer to a solution of the coupled reaction-diffusion system as a solution pair and write w = {u,v} or wi = {ui,vi} for i = 0,1,... N-1.

In general, the reaction-diffusion Eq. (15) can be understood as an abstract initial-boundary value problem: find u (t) and v(t) solutions to

dudt-A(ut)=f˜(t)anddvdt-B(v(t))=g˜(t)fora.a.t∈Iwithu0=u0,v(0)=v0andBCsonΓ1andΓ2-(16)

where Ai : V1 → V1∗ and Bi : V2 → V2∗, i = 0,1,... N-1, are independent of time t, and f˜ and g˜ independent of u and v, respectively; u0∈L2 (Ω1) and v0∈L2(Ω2).

Let us define τ > 0 a time step such that T/τ is an integer for simplicity. In the Rothe method [87,138], one has to discretise time t by backward differences and, if necessary, to approximate f˜ and g˜ at particular points t = kτ of a time grid (assumed to be equidistant for simplicity), k = 0,1,...,T/τ. This approximation can be done, for example, by convolution of f˜ and g˜ with suitable mollifiers [87,88]; however, reaction terms rising from protein-protein (protein-mRNA) interactions are smooth functions, thus well defined at each time point t.

To efficiently handle terms A and B, as they are non-linear in our reaction-diffusion systems, we can consider a certain linearisation A¯z·:V1→V1∗ of A at a point z and B¯z·:V2→V2∗ of B at a point z. Then, one can define uτk and vτk for k = 1,2,...,T/τ by

$$u\tau k-u\tau k-1\tau+A\bar{}u\tau ku\tau k-1=f\tilde{}k\tau and v\tau k-v\tau k-1\tau+B\bar{}v\tau kv\tau k-1=g\tilde{}k\tau,k=1,...,T/\tau-(17)$$

where the conditions A¯uu=Au and B¯vv=Bv are required. The initial conditions uτ0 and vτ0 are set to the original initial conditions u0 and v0 or to their suitable approximations. The Kedem-Katchalsky BCs on Γ1 and the zero-flux BCs on Γ2 for k = 1,2,...,T/τ become

$$D\partial u\tau k\partial n1=-pv\tau k-u\tau k=-D\partial v\tau k\partial n1 on\Gamma 1 and-D\partial v\tau k\partial n2=0on\Gamma 2-(18)$$

By rearranging terms in Eq. (17), one can notice that the elliptic problems in Eq. (17) with the BC (18) and the initial conditions possess unique solutions uτk∈V1 and vτk∈V2 by recalling Lax-Milgram theorem for each k = 1,...,T/τ.

The piecewise affine interpolants uτ∈[C(I;V1)]N and vτ∈[C(I;V2)]N,

$$u\tau t=tT-k-1u\tau k+k-tTu\tau k-1 for(k-1)\tau<t<k\tau,v\tau t=tT-k-1v\tau k+k-tTv\tau k-1 for(k-1)\tau<t<k\tau-(19)$$

can be used as approximations of solutions u and v to the original reaction-diffusion systems. More precisely, it can be shown that u(t) and vτ(t) weakly converge to u(t) and v(t) by limit passage for τ → 0 for a.a.t∈I. This non-rigorous procedure constructively proves existence of u(t) and v(t) in the time-space cylinders I × Ω1 and I × Ω2, more details are left to [122]. Much richer knowledge, and not only on the Rothe method, can be found in the book of Roubíček [87].

Example 3.1

For convenience and because the Leloup-Goldbeter system is less demanding on space, we can show the equations of the model (8) with the Kedem-Katchalsky BCs (10) and (11) appearing in the Rothe method. In particular, the equations in Eq. (8) may be discretised and linearised into the following elliptic problems

$$u0,\tau k-u0,\tau k-1\tau-D0\Delta u0,\tau k=VsKrKr+u1,\tau k-1r-Vmnu0,\tau kKmn+u0,\tau k-1,u1,\tau k-u1,\tau k-1\tau-D1\Delta u1,\tau k=-Vdnu1,\tau kKdn+u1,\tau k-1,v0,\tau k-v0,\tau k-1\tau-D0\Delta v0,\tau k=-Vmcv0,\tau kKmc+v0,\tau k-1,v1,\tau k-v1,\tau k-1\tau-D1\Delta v1,\tau k=ksv0,\tau k\chi tra-Vdcv1,\tau kKdc+v1,\tau k-1$$

whilst the Kedem-Katchalsky BCs (10) and (11) become

$$D0\partial u0,\tau k\partial n1=p0u0,\tau k=-D0\partial v0,\tau k\partial n1,-D1\partial u1,\tau k\partial n1=-p1v1,\tau k=-D1\partial v1,\tau k\partial n1$$

Initially, we set uτ0 = u0 and vτ0 = v0 (in our simulations u0 and v0 are positive constants in x). These linear problems yield unique, non-negative and bounded solutions u0,τk,u1,τk∈V1 and v0,τk,v1,τk∈V2 with bounded time derivatives (existence and uniqueness follows from coercivity of the problems, and thus from the Lax-Milgram theorem). It can be shown that the piecewise affine interpolants (u0,τ,u1,τ)∈[C(I;V1)]2 and (v0,τ,v1,τ)∈[C(I;V2)]2 defined in Eq. (19) are weakly convergent to (weak) solution pair

$$\{(u0,u1),(v0,v1)\} in [W11,2,2(I;V1,V1*)]2 \times [W21,2,2(I;V2,V2*)]2$$

It is, however, worth mentioning that the computational time needed to obtain a solution to the Leloup-Goldbeter model by the semi-implicit Rothe method is 3.3-fold lower than the time required to solve the same problem with the Newton method (computations with T = 72 h and τ = 0.1 took 105 s in the semi-implicit Rothe method implementation and 344 s in the Newton method with the accuracy tolerance tol = 10-3 needed to terminate iterations). Similarly, the Rothe method is 5.2-faster than the Newton method when computing solution to the p53 model (within t= 400 time units and τ= 0.2 took 37 min in the semi-implicit Rothe method implementation and 193 min in the Newton method with tol =10-3).

For the Leloup-Goldbeter and p53 reaction-diffusion systems presented in this article, the semi-implicit Rothe method and the Newton method have been implemented in the freely available solver FreeFem++ [63] and we ran 2D simulations on a disk-shaped cell with radius 10μm, 2.4 GHz Intel Core icontain contain 7 processor and 8 GB (1600 MHz) memory.

High accuracy spline explicit group (seg) approximations

Consider the two-dimensional elliptic partial differential equation

$$\partial 2u\partial x2+\partial 2u\partial y2=D(x)\partial u\partial x+g(x,y),(x,y)\in\Omega-(21)$$

which is defined in the solution domain Ω = {(x, y):0 < x, y < 1}, where functions D(x) and g(x, y) ∈ C2(Ω). The corresponding Dirichlet boundary conditions are given by

$$u(x, y) = \psi(x, y), (x, y) \in \partial\Omega\text{-}(22)$$

where $\partial\Omega$ is its boundary. These types of problems arise very frequently in different areas of applied mathematics and physics such as convection-diffusion equation which describes the transport phenomena, and the Poisson's equation which is broadly used in electrostatics, mechanical engineering and theoretical physics. Thus, solving elliptic differential equation has been of interest to many authors [143-145]. In 1968, Bickley [146] suggested the use of cubic splines for solving a linear ordinary differential. Following this, Albasiny and Hoskins [147] approximated the solutions by applying the cubic spline interpolation introduced by Ahlberg et al. [152], which leads to a matrix system of tri-diagonal instead of upper Hessenberg form which was obtained by Bickley [146]. The cubic spline method suggested by Bickley [146] was then examined by Fyfe [153]. Fyfe concluded that spline method is better than the usual finite difference method in terms of its accuracy and also its flexibility to get the approximation at any point in the domain. Due to its simplicity, many researchers started to work on spline methods for solving boundary value problems [154-157]. To mention a few, Bialecki et al. [158] formulated a new fourth order one step nodal bicubic spline collocation method for the solution of various elliptic boundary value problems. Mohanty and Gopal [159] proposed a high accuracy cubic spline finite difference approximation of O(k2 + h4) accuracy for the solution of non-linear wave equation. Goh et al. [160] discussed the solution for one-dimensional heat and advection-diffusion by using a combination of finite difference approach and cubic B-spline method. The numerical results show that the EG method is simpler to program compared to the block (line) iterative methods and it requires less storage. However, this method was solely formulated using the usual standard finite difference discretization which restricts the solutions at only certain points of the solution domain. This, thus, motivate us to adopt the idea in using splines in the formulation of the group methods [161,162]. In this paper, a new method, namely spline explicit group (SEG) iterative method, which incorporates cubic spline with group iterative scheme, is developed for solving the elliptic problems. Using a cubic spline approximation in the x-direction and central difference in the y-direction, we obtain a new three level implicit nine-point compact finite difference formulation [163,164]. Then, a four point explicit group iterative scheme is applied to the obtained system. The performance of the method will be investigated via two benchmark problems, that is the convection-diffusion equation and Poisson's equation [165,166].

The cubic spline approximation and numerical scheme

Here, the solution domain $\Omega = [0,1] \times [0,1]$ is divided by $h > 0$ in x-direction and $k > 0$ in y-direction. Therefore, the grid points (xl, ym) are represented as xl = lh and ym = mk, l = 0, 1, …, Nx, m = 0, 1, …, Ny, where Nx and Ny are positive integers. Let Ul, m be the approximation solution of ul, m at the grid point (xl, ym).

Suppose that Sm(x) is the m-th mesh row cubic spline polynomial which interpolates the value Ul, m at (xl, ym), is given by [152]

$$Sm(x)=(xl-x)36hMl-1,m+(x-xl-1)36hMl,m+(xl-xh)(Ul-1,m-h26Ml-1,m)+(x-xl-1h)(Ul,m-h26Ml,m)\text{-}(23)$$

for $xl-1 \le x \le xl$, where l = 1, 2, …, Nx and m = 0, 1, 2, …, Ny. For each m-th mesh row, the cubic spline Sm(x) satisfies the following properties

Sm(x) coincides with a polynomial of degree three on each [xl-1, xl], l = 1, 2, …, Nx, m = 0, 1, 2, …, Ny

$Sm(x) \in C2[0, 1]$, and

Sm(xl) = Ul, m, l = 0, 1, 2, …, Nx, m = 0, 1, 2, …, Ny

The derivatives of cubic spline Sm(x) can be obtained as below

$$Sm'(x)=-(xl-x)22hMl-1,m+(x-xl-1)22hMl,m+Ul,m-Ul-1,mh-h6[Ml,m-Ml-1,m]\text{-}(24)$$

$$Sm''(x)=(xl-x)hMl-1,m+(x-xl-1)hMl,m\text{-}(25)$$

And, from Eq (1), it gives

$$Ml,m=Sm''(xl)=Uxxl,m=-Uyyl,m+DlUxl,m+gl,m\text{-}(26)$$

When x = xl, Eq (4) becomes

$$Sm'(xl)=Uxl,m=Ul,m-Ul-1,mh+h6[Ml-1,m+2Ml,m]\text{-}(27)$$

Similarly, for $x \in [xl, xl+1]$, it gives

$$Sm'(xl)=Uxl,m=Ul+1,m-Ul,mh-h6[Ml+1,m+2Ml,m]\text{-}(28)$$

Combining both Eqs (7) and (8), the following approximation can be obtained

$$Sm'(xl)=Uxl,m=Ul+1,m-Ul-1,m2h-h12[Ml+1,m-Ml-1,m]\text{-}(29)$$

Further, we have

$$Sm'(xl+1)=Uxl+1,m=Ul+1,m-Ul,mh+h6[Ml,m+2Ml+1,m]\text{-}(210)$$

$$Sm'(xl-1)=Uxl-1,m=Ul,m-Ul-1,mh-h6[Ml,m+2Ml-1,m]\text{-}(211)$$

By using the continuity of first derivative at (xl, ym), which is, Sm′(xl+)=Sm′(xl-), the following relation can be obtained

Ul+1,m-2Ul,m+Ul-1,m=h26(Ml+1,m+4Ml,m+Ml-1,m)-(212)

The following approximations are considered

U¯yyl,m=(Ul,m+1-2Ul,m+Ul,m-1)/k2-(213a)

U¯yyl+1,m=(Ul+1,m+1-2Ul+1,m+Ul+1,m-1)/k2-(213b)

U¯yyl-1,m=(Ul-1,m+1-2Ul-1,m+Ul-1,m-1)/k2-(213c)

U¯xl,m=(Ul+1,m-Ul-1,m)/(2h)-(214a)

U¯xl+1,m=(3Ul+1,m-4Ul,m+Ul-1,m)/(2h)-(214b)

U¯xl-1,m=(-3Ul-1,m+4Ul,m-Ul+1,m)/(2h)-(214c)

Eqs (14b) and (14c) are obtained from the second-order one-sided finite difference scheme. For the derivatives of Sm(x), we consider

M¯l,m=-U¯yyl,m+DlU¯xl,m+gl,m-(215a)

M¯l+1,m=-U¯yyl+1,m+Dl+1U¯xl+1,m+gl+1,m-(215b)

M¯l-1,m=-U¯yyl-1,m+Dl-1U¯xl-1,m+gl-1,m-(215c)

U¯¯xl+1,m=Ul+1,m-Ul,mh+h6[M¯l,m+2M¯l+1,m]-(215d)

U¯¯xl-1,m=Ul,m-Ul-1,mh-h6[M¯l,m+2M¯l-1,m]-(215e)

U^xl,m=Ul+1,m-Ul-1,m2h-h12[M¯l+1,m-M¯l-1,m]-(215f)

By using Taylor series expansion about the grid point (xl, ym), Eq (1) can be written as

(Ul+1,m-2Ul,m+Ul-1,m)+h212[U¯yyl+1,m+U¯yyl-1,m+10U¯yyl,m]=h212[Dl+1U¯¯xl+1,m+Dl-1U¯¯xl-1,m+10DlU^xl,m]+h212[gl+1,m+gl-1,m+10gl,m]+Tl,m-(216)

where Tl, m is the local truncation error. Substituting the above approximations (13)-(15) into (16), it results

{-2+12λ2-λ2h(Dl+1+5Dl)+λ2h212[(5Dl-2Dl-1)Dl-1-3(2Dl+1-5Dl)Dl+1-(Dl+1-Dl-1)Dl]-h3(2Dl+1-5Dl)}Ul+1,m+{-20-24λ2-λ2h(Dl-1-Dl+1)+λ2h23[(2Dl+1-5Dl)Dl+1-(5Dl-2Dl-1)Dl-1]-h3(Dl+1-Dl-1)}Ul,m+{-2+12λ2+λ2h(5Dl+Dl-1)+λ2h212[3(5Dl-2Dl-1)Dl-1-(2Dl+1-5Dl)Dl+1+(Dl+1-Dl-1)Dl]-h3(5Dl-2Dl-1)}Ul-1,m+[1+h6(2Dl+1-5Dl)](Ul+1,m+1+Ul+1,m-1)+[1+h6(5Dl-2Dl-1)](Ul-1,m+1+Ul-1,m-1)+[10+h6(Dl+1-Dl-1)](Ul,m+1+Ul,m-1)=k2{[1+h6(2Dl+1-5Dl)]gl+1,m+[1+h6(5Dl-2Dl-1)]gl-1,m+[10+h6(Dl+1-Dl-1)]gl,m}+Tl,m-(217)

where λ is the mesh ratio, denoted by λ = (k/h). If the singular terms like 1x appear in the functions D(x) and/or g(x, y), which is unable to evaluate at x = 0. The following approximations are considered

Dl±1=D00±hD10+h22D20±O(h3)-(218a)

gl±1,m=g00±hg10+h22g20±O(h3)-(218b)

where

Wab=∂a+bW(xl,ym)∂xa∂yb,W=Dandg

Thus, neglecting the higher order terms and local truncation error, Eq (17) can be written as

{-2+12λ2-12λ2h(12D00+h2D20)-λ2h2(D10-D00D00)+hD00}Ul+1,m+{-20-24λ2+2λ2h2(D10-D00D00)-23h2D10}Ul,m+{-2+12λ2+12λ2h(12D00+h2D20)+λ2h2(D00D00-D10)-hD00}Ul-1,m+(1-h2D00)(Ul+1,m+1+Ul+1,m-1)+(1+h2D00)(Ul-1,m+1+Ul-1,m-1)+(10+h23D10)(Ul,m+1+Ul,m-1)=k2[12g00+h2(g20+D10g00-D00g10)]=Gl,m-(219)

This modified equation retains its order of accuracy everywhere throughout the solution region, moreover in the vicinity of the singularity. Note that, this proposed scheme

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = D(x)\frac{\partial u}{\partial x} + g(x, y), (x, y) \in \Omega$$

$$S_m(x) = \frac{(x_l - x)^3}{6h} M_{l-1,m} + \frac{(x - x_{l-1})^3}{6h} M_{l,m} + \left(\frac{x_l - x}{h}\right)\left(U_{l-1,m} - \frac{h^2}{6} M_{l-1,m}\right) + \left(\frac{x - x_{l-1}}{h}\right)\left(U_{l,m} - \frac{h^2}{6} M_{l,m}\right)$$

$$S_m^{'}(x) = \frac{-(x_l - x)^2}{2h} M_{l-1,m} + \frac{(x - x_{l-1})^2}{2h} M_{l,m} + \frac{U_{l,m} - U_{l-1,m}}{h} - \frac{h}{6}\left[M_{l,m} - M_{l-1,m}\right]$$

$$S_m^{''}(x) = \frac{(x_l - x)}{h} M_{l-1,m} + \frac{(x - x_{l-1})}{h} M_{l,m}$$

$$M_{l,m} = S_m^{''}(x_l) = U_{xxl,m} = -U_{yyl,m} + D_l U_{xl,m} + g_{l,m}$$

$$S_m^{'}(x_l) = U_{xl,m} = \frac{U_{l,m} - U_{l-1,m}}{h} + \frac{h}{6}\left[M_{l-1,m} + 2M_{l,m}\right]$$

$$S_m^{'}(x_l) = U_{xl,m} = \frac{U_{l+1,m} - U_{l,m}}{h} - \frac{h}{6}\left[M_{l+1,m} + 2M_{l,m}\right]$$

$$S_m^{'}(x_l) = U_{xl,m} = \frac{U_{l+1,m} - U_{l-1,m}}{2h} - \frac{h}{12}\left[M_{l+1,m} - M_{l-1,m}\right]$$

$$S_m^{'}(x_{l+1}) = U_{xl+1,m} = \frac{U_{l+1,m} - U_{l,m}}{h} + \frac{h}{6}\left[M_{l,m} + 2M_{l+1,m}\right]$$

$$S_m^{'}(x_{l-1}) = U_{xl-1,m} = \frac{U_{l,m} - U_{l-1,m}}{h} - \frac{h}{6}\left[M_{l,m} + 2M_{l-1,m}\right]$$

$$S_m^{'}\left(x_l^{+}\right) = S_m^{'}\left(x_l^{-}\right)$$

$$U_{l+1,m} - 2U_{l,m} + U_{l-1,m} = \frac{h^2}{6}\left(M_{l+1,m} + 4M_{l,m} + M_{l-1,m}\right)$$

$$\overline{U}_{yyl,m} = (U_{l,m+1} - 2U_{l,m} + U_{l,m-1})/k^2$$

$$\overline{U}_{yyl+1,m} = (U_{l+1,m+1} - 2U_{l+1,m} + U_{l+1,m-1})/k^2$$

$$\overline{U}_{xl,m} = (U_{l+1,m} - U_{l-1,m})/(2h)$$

$$\overline{U}_{xl+1,m} = (3U_{l+1,m} - 4U_{l,m} + U_{l-1,m})/(2h)$$

$$\overline{U}_{xl-1,m} = (-3U_{l-1,m} + 4U_{l,m} - U_{l+1,m})/(2h)$$

$$\overline{M}_{l,m} = -\overline{U}_{yyl,m} + D_l \overline{U}_{xl,m} + g_{l,m}$$

$$\overline{M}_{l+1,m} = -\overline{U}_{yyl+1,m} + D_{l+1} \overline{U}_{xl+1,m} + g_{l+1,m}$$

$$\overline{M}_{l-1,m} = -\overline{U}_{yyl-1,m} + D_{l-1} \overline{U}_{xl-1,m} + g_{l-1,m}$$

$$\overline{\overline{U}}_{xl+1,m} = \frac{U_{l+1,m} - U_{l,m}}{h} + \frac{h}{6}\left[\overline{M}_{l,m} + 2\overline{M}_{l+1,m}\right]$$

$$\overline{\overline{U}}_{xl-1,m} = \frac{U_{l,m} - U_{l-1,m}}{h} - \frac{h}{6}\left[\overline{M}_{l,m} + 2\overline{M}_{l-1,m}\right]$$

$$\widehat{U}_{xl,m} = \frac{U_{l+1,m} - U_{l-1,m}}{2h} - \frac{h}{12}\left[\overline{M}_{l+1,m} - \overline{M}_{l-1,m}\right]$$

$$\left(U_{l+1,m} - 2U_{l,m} + U_{l-1,m}\right) + \frac{h^2}{12}\left[\overline{U}_{yyl+1,m} + \overline{U}_{yyl-1,m} + 10\overline{U}_{yyl,m}\right]$$

$$= \frac{h^2}{12}\left[D_{l+1}\overline{\overline{U}}_{xl+1,m} + D_{l-1}\overline{\overline{U}}_{xl-1,m} + 10D_l\widehat{U}_{xl,m}\right] + \frac{h^2}{12}\left[g_{l+1,m} + g_{l-1,m} + 10g_{l,m}\right] + T_{l,m}$$

$$\left\{-2 + 12\lambda^2 - \lambda^2 h(D_{l+1} + 5D_l) + \frac{\lambda^2 h^2}{12}\left[(5D_l - 2D_{l-1})D_{l-1} - 3(2D_{l+1} - 5D_l)D_{l+1} - (D_{l+1} - D_{l-1})D_l\right] - \frac{h}{3}(2D_{l+1} - 5D_l)\right\} U_{l+1,m}$$

$$+\left\{-20 - 24\lambda^2 - \lambda^2 h(D_{l-1} - D_{l+1}) + \frac{\lambda^2 h^2}{3}\left[(2D_{l+1} - 5D_l)D_{l+1} - (5D_l - 2D_{l-1})D_{l-1}\right] - \frac{h}{3}(D_{l+1} - D_{l-1})\right\} U_{l,m}$$

$$+\left\{-2 + 12\lambda^2 + \lambda^2 h(5D_l + D_{l-1}) + \frac{\lambda^2 h^2}{12}\left[3(5D_l - 2D_{l-1})D_{l-1} - (2D_{l+1} - 5D_l)D_{l+1} + (D_{l+1} - D_{l-1})D_l\right] - \frac{h}{3}(5D_l - 2D_{l-1})\right\} U_{l-1,m}$$

$$+\left[1 + \frac{h}{6}(2D_{l+1} - 5D_l)\right](U_{l+1,m+1} + U_{l+1,m-1})$$

$$+\left[1 + \frac{h}{6}(5D_l - 2D_{l-1})\right](U_{l-1,m+1} + U_{l-1,m-1})$$

$$+\left[10 + \frac{h}{6}(D_{l+1} - D_{l-1})\right](U_{l,m+1} + U_{l,m-1})$$

$$k^2\left\{\left[1 + \frac{h}{6}(2D_{l+1} - 5D_l)\right]g_{l+1,m} + \left[1 + \frac{h}{6}(5D_l - 2D_{l-1})\right]g_{l-1,m} + \left[10 + \frac{h}{6}(D_{l+1} - D_{l-1})\right]g_{l,m}\right\} + T_{l,m}$$

$$D_{l\pm1} = D_{00} \pm hD_{10} + \tfrac{h^2}{2} D_{20} \pm O(h^3)$$

$$g_{l\pm1,m} = g_{00} \pm hg_{10} + \tfrac{h^2}{2} g_{20} \pm O(h^3)$$

$$W_{ab} = \frac{\partial^{a+b} W(x_l, y_m)}{\partial x^a \partial y^b},\ W = D \text{and} g$$

$$\{-2 + 12\lambda^2 - \tfrac{1}{2}\lambda^2 h(12D_{00} + h^2 D_{20}) - \lambda^2 h^2 (D_{10} - D_{00}D_{00}) + hD_{00}\} U_{l+1,m}$$

$$+ \quad \{-20 - 24\lambda^2 + 2\lambda^2 h^2 (D_{10} - D_{00}D_{00}) - \tfrac{2}{3} h^2 D_{10}\} U_{l,m}$$

$$+ \quad \{-2 + 12\lambda^2 + \tfrac{1}{2}\lambda^2 h(12D_{00} + h^2 D_{20}) + \lambda^2 h^2 (D_{00}D_{00} - D_{10}) - hD_{00}\} U_{l-1,m}$$

$$+ \quad (1 - \tfrac{h}{2} D_{00})(U_{l+1,m+1} + U_{l+1,m-1}) + (1 + \tfrac{h}{2} D_{00})(U_{l-1,m+1} + U_{l-1,m-1})$$

$$+ \quad \left(10 + \tfrac{h^2}{3} D_{10}\right)(U_{l,m+1} + U_{l,m-1})$$

$$= \quad k^2 [12g_{00} + h^2 (g_{20} + D_{10}g_{00} - D_{00}g_{10})] = G_{l,m}$$

$$\begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ a_5 & a_1 & a_4 & a_6 \\ a_6 & a_4 & a_1 & a_5 \\ a_4 & a_3 & a_2 & a_1 \end{bmatrix} \begin{bmatrix} U_{l,m} \\ U_{l+1,m} \\ U_{l+1,m+1} \\ U_{l,m+1} \end{bmatrix} = \begin{bmatrix} rhs_{l,m} \\ rhs_{l+1,m} \\ rhs_{l+1,m+1} \\ rhs_{l,m+1} \end{bmatrix}$$

$$a_1 = -20 - 24\lambda^2 + 2\lambda^2 h^2 (D_{10} - D_{00}D_{00}) - \tfrac{2}{3} h^2 D_{10}$$

$$a_2 = -2 + 12\lambda^2 - \tfrac{1}{2}\lambda^2 h(12D_{00} + h^2 D_{20}) - \lambda^2 h^2 (D_{10} - D_{00}D_{00}) + hD_{00}$$

$$a_3 = 1 - \tfrac{h}{2} D_{00}$$

$$a_4 = 10 + \tfrac{h^2}{3} D_{10}$$

$$a_5 = -2 + 12\lambda^2 + \tfrac{1}{2}\lambda^2 h(12D_{00} + h^2 D_{20}) + \lambda^2 h^2 (D_{00}D_{00} - D_{10}) - hD_{00}$$

$$a_6 = 1 + \tfrac{h}{2} D_{00}$$

$$b_1 = a_1^3 - a_1 a_4^2 - a_1 a_2 a_5 + a_3 a_4 a_5 - a_1 a_3 a_6 + a_2 a_4 a_6$$

$$b_2 = -a_1^2 a_2 + 2a_1 a_3 a_4 - a_2 a_4^2 + a_2^2 a_5 - a_3^2 a_5$$

$$b_3 = -a_1^2 a_3 + 2a_1 a_2 a_4 - a_3 a_4^2 - a_2^2 a_6 + a_3^2 a_6$$

$$b_4 = -a_1^2 a_4 + a_4^3 + a_1 a_3 a_5 - a_2 a_4 a_5 + a_1 a_2 a_6 - a_3 a_4 a_6$$

$$b_5 = -a_1^2 a_5 - a_4^2 a_5 + a_2 a_5^2 + 2a_1 a_4 a_6 - a_2 a_6^2$$

$$b_6 = 2a_1 a_4 a_5 - a_3 a_5^2 - a_1^2 a_6 - a_4^2 a_6 + a_3 a_6^2$$

$$\left(U_{l+1,m} - 2U_{l,m} + U_{l-1,m}\right) + \tfrac{h^2}{12}\left[\overline{U}_{yyl+1,m} + \overline{U}_{yyl-1,m} + 10\overline{U}_{yyl,m}\right]$$

$$= \quad \tfrac{h^2}{12}\left[D_{l+1}\overline{\overline{U}}_{xl+1,m} + D_{l-1}\overline{\overline{U}}_{xl-1,m} + 10D_l \widehat{U}_{xl,m}\right] + \tfrac{h^2}{12}\left[g_{l+1,m} + g_{l-1,m} + 10g_{l,m}\right] + T_{l,m}$$

$$\left\{-2 + 12\lambda^2 - \lambda^2 h(D_{l+1} + 5D_l) + \tfrac{\lambda^2 h^2}{12}\left[(5D_l - 2D_{l-1})D_{l-1} - 3(2D_{l+1} - 5D_l)D_{l+1}\right.\right.$$

$$\left.\left. -(D_{l+1} - D_{l-1})D_l\right] - \tfrac{h}{3}(2D_{l+1} - 5D_l)\right\} U_{l+1,m}$$

$$+\left\{-20 - 24\lambda^2 - \lambda^2 h(D_{l-1} - D_{l+1}) + \tfrac{\lambda^2 h^2}{3}\left[(2D_{l+1} - 5D_l)D_{l+1} - (5D_l - 2D_{l-1})D_{l-1}\right]\right.$$

$$\left. -\tfrac{h}{3}(D_{l+1} - D_{l-1})\right\} U_{l,m}$$

$$+\left\{-2 + 12\lambda^2 + \lambda^2 h(5D_l + D_{l-1}) + \tfrac{\lambda^2 h^2}{12}\left[3(5D_l - 2D_{l-1})D_{l-1} - (2D_{l+1} - 5D_l)D_{l+1}\right.\right.$$

$$\left.\left. +(D_{l+1} - D_{l-1})D_l\right] - \tfrac{h}{3}(5D_l - 2D_{l-1})\right\} U_{l-1,m}$$

$$+\left[1 + \tfrac{h}{6}(2D_{l+1} - 5D_l)\right](U_{l+1,m+1} + U_{l+1,m-1})$$

$$+\left[1 + \tfrac{h}{6}(5D_l - 2D_{l-1})\right](U_{l-1,m+1} + U_{l-1,m-1})$$

$$+\left[10 + \tfrac{h}{6}(D_{l+1} - D_{l-1})\right](U_{l,m+1} + U_{l,m-1})$$

$$k^2\left\{\left[1 + \tfrac{h}{6}(2D_{l+1} - 5D_l)\right]g_{l+1,m} + \left[1 + \tfrac{h}{6}(5D_l - 2D_{l-1})\right]g_{l-1,m}\right.$$

$$+\left[10+\frac{h}{6}\left(D_{l+1}-D_{l-1}\right)\right]g_{l,m}\Big\}+T_{l,m}$$

$$\frac{1}{x}$$

$$D_{l\pm1}=D_{00}\pm hD_{10}+\frac{h^2}{2}D_{20}\pm O(h^3)$$

$$g_{l\pm1,m}=g_{00}\pm hg_{10}+\frac{h^2}{2}g_{20}\pm O(h^3)$$

$$W_{ab}=\frac{\partial^{a+b}W(x_l,y_m)}{\partial x^a\partial y^b},W=D\text{and}g$$

$$\left\{-2+12\lambda^2-\tfrac{1}{2}\lambda^2h(12D_{00}+h^2D_{20})-\lambda^2h^2(D_{10}-D_{00}D_{00})+hD_{00}\right\}U_{l+1,m}$$

$$+\quad\left\{-20-24\lambda^2+2\lambda^2h^2(D_{10}-D_{00}D_{00})-\tfrac{2}{3}h^2D_{10}\right\}U_{l,m}$$

$$+\quad\left\{-2+12\lambda^2+\tfrac{1}{2}\lambda^2h(12D_{00}+h^2D_{20})+\lambda^2h^2(D_{00}D_{00}-D_{10})-hD_{00}\right\}U_{l-1,m}$$

$$+\quad\left(1-\tfrac{h}{2}D_{00}\right)\left(U_{l+1,m+1}+U_{l+1,m-1}\right)+\left(1+\tfrac{h}{2}D_{00}\right)\left(U_{l-1,m+1}+U_{l-1,m-1}\right)$$

$$+\quad\left(10+\frac{h^2}{3}D_{10}\right)\left(U_{l,m+1}+U_{l,m-1}\right)$$

$$=\quad k^2\left[12g_{00}+h^2\left(g_{20}+D_{10}g_{00}-D_{00}g_{10}\right)\right]=G_{l,m}$$

$$\begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ a_5 & a_1 & a_4 & a_6 \\ a_6 & a_4 & a_1 & a_5 \\ a_4 & a_3 & a_2 & a_1 \end{bmatrix} \begin{bmatrix} U_{l,m} \\ U_{l+1,m} \\ U_{l+1,m+1} \\ U_{l,m+1} \end{bmatrix} = \begin{bmatrix} rhs_{l,m} \\ rhs_{l+1,m} \\ rhs_{l+1,m+1} \\ rhs_{l,m+1} \end{bmatrix}$$

$$a_1 = -20 - 24\lambda^2 + 2\lambda^2 h^2 (D_{10} - D_{00}D_{00}) - \tfrac{2}{3} h^2 D_{10}$$

$$a_2 = -2 + 12\lambda^2 - \tfrac{1}{2}\lambda^2 h(12D_{00} + h^2 D_{20}) - \lambda^2 h^2 (D_{10} - D_{00}D_{00}) + hD_{00}$$

$$a_3 = 1 - \tfrac{h}{2} D_{00}$$

$$a_4 = 10 + \tfrac{h^2}{3} D_{10}$$

$$a_5 = -2 + 12\lambda^2 + \tfrac{1}{2}\lambda^2 h(12D_{00} + h^2 D_{20}) + \lambda^2 h^2 (D_{00}D_{00} - D_{10}) - hD_{00}$$

$$a_6 = 1 + \tfrac{h}{2} D_{00}$$

$$rhs_{l,m} = -a_5 U_{l-1,m} - a_3 U_{l+1,m-1} - a_6 (U_{l-1,m+1} + U_{l-1,m-1}) - a_4 U_{l,m-1} + G_{l,m}$$

$$rhs_{l+1,m} = -a_2 U_{l+2,m} - a_3 (U_{l+2,m+1} + U_{l+2,m-1}) - a_6 U_{l,m-1} - a_4 U_{l+1,m-1} + G_{l+1,m}$$

$$rhs_{l+1,m+1} = -a_2 U_{l+2,m+1} - a_3 (U_{l+2,m+2} + U_{l+2,m}) - a_6 U_{l,m+2} - a_4 U_{l+1,m+2} + G_{l+1,m+1}$$

$$rhs_{l,m+1} = -a_5 U_{l-1,m+1} - a_3 U_{l+1,m+2} - a_6 (U_{l-1,m+2} + U_{l-1,m}) - a_4 U_{l,m+2} + G_{l,m+1}$$

$$\begin{bmatrix} U_{l,m} \\ U_{l+1,m} \\ U_{l+1,m+1} \\ U_{l,m+1} \end{bmatrix} = \frac{1}{denom} \begin{bmatrix} b_1 & b_2 & b_3 & b_4 \\ b_5 & b_1 & b_4 & b_6 \\ b_6 & b_4 & b_1 & b_5 \\ b_4 & b_3 & b_2 & b_1 \end{bmatrix} \begin{bmatrix} rhs_{l,m} \\ rhs_{l+1,m} \\ rhs_{l+1,m+1} \\ rhs_{l,m+1} \end{bmatrix}$$

$$denom = a_1^4 - 2a_1^2 a_4^2 + a_4^4 - 2a_1^2 a_2 a_5 + 4a_1 a_3 a_4 a_5 - 2a_2 a_4^2 a_5 + a_2^2 a_5^2 - a_3^2 a_5^2 - 2a_1^2 a_3 a_6$$
$$+ 4a_1 a_2 a_4 a_6 - 2a_3 a_4^2 a_6 - a_2^2 a_6^2 + a_3^2 a_6^2$$

$$b_1 = a_1^3 - a_1 a_4^2 - a_1 a_2 a_5 + a_3 a_4 a_5 - a_1 a_3 a_6 + a_2 a_4 a_6$$

$$b_2 = -a_1^2 a_2 + 2a_1 a_3 a_4 - a_2 a_4^2 + a_2^2 a_5 - a_3^2 a_5$$

$$b_3 = -a_1^2 a_3 + 2a_1 a_2 a_4 - a_3 a_4^2 - a_2^2 a_6 + a_3^2 a_6$$

$$b_4 = -a_1^2 a_4 + a_4^3 + a_1 a_3 a_5 - a_2 a_4 a_5 + a_1 a_2 a_6 - a_3 a_4 a_6$$

$$b_5 = -a_1^2 a_5 - a_4^2 a_5 + a_2 a_5^2 + 2a_1 a_4 a_6 - a_2 a_6^2$$

$$b_6 = 2a_1 a_4 a_5 - a_3 a_5^2 - a_1^2 a_6 - a_4^2 a_6 + a_3 a_6^2$$

$$A = \begin{bmatrix} D & V & \\ L & D & V \\ & L & D \end{bmatrix}$$

$$D = \begin{bmatrix} R_0 & R_2 & \\ R_5 & R_0 & R_2 \\ & R_5 & R_0 \end{bmatrix}, V = \begin{bmatrix} R_4 & R_3 & \\ R_6 & R_4 & R_3 \\ & R_6 & R_4 \end{bmatrix}, L = \begin{bmatrix} R_4^{'} & R_3^{'} & \\ R_6^{'} & R_4^{'} & R_3^{'} \\ & R_6^{'} & R_4^{'} \end{bmatrix}$$

$$R_0 = \begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ a_5 & a_1 & a_4 & a_6 \\ a_6 & a_4 & a_1 & a_5 \\ a_4 & a_3 & a_2 & a_1 \end{bmatrix}, R_2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ a_2 & 0 & 0 & a_3 \\ a_3 & 0 & 0 & a_2 \\ 0 & 0 & 0 & 0 \end{bmatrix}, R_3 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ a_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$R_4 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ a_6 & a_4 & 0 & 0 \\ a_4 & a_3 & 0 & 0 \end{bmatrix}, R_5 = \begin{bmatrix} 0 & a_5 & a_6 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & a_6 & a_5 & 0 \end{bmatrix}, R_6 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & a_6 & 0 & 0 \end{bmatrix},$$

$$R_3^{'} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & a_3 \\ 0 & 0 & 0 & 0 \end{bmatrix}, R_4^{'} = \begin{bmatrix} 0 & 0 & a_3 & a_4 \\ 0 & 0 & a_4 & a_6 \\ 0 & 0 & 0 & 0 \end{bmatrix}, R_6^{'} = \begin{bmatrix} 0 & 0 & a_6 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$A^E = diag\{R_0^{-1}\}A$$

$$b^E = diag\{R_0^{-1}\}b$$

$$R_j^{'}$$

$$R_0^{-1}R_i$$

$$R_0^{-1}R_j^{'}$$

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = \beta \frac{\partial u}{\partial x}, 0 < x, y < 1$$

$$u(x, y) = e^{\frac{\beta x}{2}} \frac{\sin \pi y}{\sinh \sigma} \left[2e^{\frac{-\beta}{2}} \sinh \sigma x + \sinh \sigma(1 - x) \right]$$

$$\sigma^2 = \pi^2 + \frac{\beta^2}{4}$$

$$\frac{\partial^2 u}{\partial r^2} + \frac{\partial^2 u}{\partial z^2} + \frac{1}{r}\frac{\partial u}{\partial r} = \cosh z\left(5r \cosh r + 2\left(2 + r^2\right) \sinh r\right), 0 < r, z < 1$$

$$D(r) = -\frac{1}{r}$$

High Accuracy Spline Explicit Group (SEG) Approximation for Two Dimensional Elliptic Boundary Value Problems

(219) is of O(k2 + k2h2 + h4) and applicable to both singular and non-singular elliptic equations of the form (1).

Spline explicit group method

We apply Eq (19) to any group of four points on the solution domain. Then, a (4 × 4) system as below, can be obtained

[a1a2a3a4a5a1a4a6a6a4a1a5a4a3a2a1][Ul,mUl+1,mUl+1,m+1Ul,m+1]=[rhsl,mrhsl+1,mrhsl+1,m+1rhsl,m+1]-(220)

where

a1=-20-24λ2+2λ2h2(D10-D00D00)-23h2D10a2=-2+12λ2-12λ2h(12D00+h2D20)-λ2h2(D10-D00D00)+hD00a3=1-h2D00a4=10+h23D10a5=-2+12λ2+12λ2h(12D00+h2D20)+λ2h2(D00D00-D10)-hD00a6=1+h2D00

and

rhsl,m=-a5Ul-1,m-a3Ul+1,m-1-a6(Ul-1,m+1+Ul-1,m-1)-a4Ul,m-1+Gl,mrhsl+1,m=-a2Ul+2,m-a3(Ul+2,m+1+Ul+2,m-1)-a6Ul,m-1-a4Ul+1,m-1+Gl+1,mrhsl+1,m+1=-a2Ul+2,m+1-a3(Ul+2,m+2+Ul+2,m)-a6Ul,m+2-a4Ul+1,m+2+Gl+1,m+1rhsl,m+1=-a5Ul-1,m+1-a3Ul+1,m+2-a6(Ul-1,m+2+Ul-1,m)-a4Ul,m+2+Gl,m+1

Eq (20) can be inverted and written in explicit forms

[Ul,mUl+1,mUl+1,m+1Ul,m+1]=1denom[b1b2b3b4b5b1b4b6b6b4b1b5b4b3b2b1][rhsl,mrhsl+1,mrhsl+1,m+1rhsl,m+1]-(221)

where

denom=a14-2a12a42+a44-2a12a2a5+4a1a3a4a5-2a2a42a5+a22a52-a32a52-2a12a3a6+4a1a2a4a6-2a3a42a6-a22a62+a32a62

and

b1=a13-a1a42-a1a2a5+a3a4a5-a1a3a6+a2a4a6b2=-a12a2+2a1a3a4-a2a42+a22a5-a32a5b3=-a12a3+2a1a2a4-a3a42-a22a6+a32a6b4=-a12a4+a43+a1a3a5-a2a4a5+a1a2a6-a3a4a6b5=-a12a5-a42a5+a2a52+2a1a4a6-a2a62b6=2a1a4a5-a3a52-a12a6-a42a6+a3a62

The Gauss-Seidel technique is employed to accelerate the convergence process. Iterations are generated in groups of four points over the entire spatial domain until the convergence test is satisfied [167]. Once the approximations Ul,m had been calculated, the value of Ml,m can be easily obtained by solving the system generated by (12). Then, the piecewise polynomial of the function can be obtained from Eq (3). Finally, the approximate solution at any point at m-th mesh row can be easily calculated.

Computational Molecule of Eq. (20)

Applying Eq (21) to each of the group in natural row ordering will lead to a linear system

AU = b

where the matrix of coefficient A is given by

A=[DVLDVLD]-(222)

with

D=[R0R2R5R0R2R5R0],V=[R4R3R6R4R3R6R4],L=[R4′R3′R6′R4′R3′R6′R4′]

The submatrices are given by

R0=[a1a2a3a4a5a1a4a6a6a4a1a5a4a3a2a1],R2=[0000a200a3a300a20000],R3=[00000000a30000000],

R4=[00000000a6a400a4a300],R5=[0a5a60000000000a6a50],R6=[0000000000000a600],

R3′=[0000000a300000000],R4′=[00a3a400a4a600000000],R6′=[00a600000000000000]

In order to derive the explicit formulae, the matrix A is transformed into AE and vector b is modified into bE, where,

AE=diag {R0-1}AbE=diag{R0-1}b

The block structure of AE is the same as matrix A with the nonzero block R0 replaced by identity matrices, I and the blocks Ri and Rj′, replaced by R0-1Ri, i = 0, 2, 3, 4, 5, 6 and R0-1Rj′, j = 3,4,6 respectively. Since the coefficient matrix [168] is block tridiagonal with non-vanishing diagonal element, it is π-consistently ordered and has property-A (π) [168]. Thus, the theory of block S.O.R. is also applicable to the SEG iterative method and therefore, is convergent.

Computational complexity analysis

In order to show the efficiency of the proposed method, computational complexity of the SEG iterative method is examined [169]. Assume that the solution domain is discretized into even intervals, Nx and Ny in x- and y-directions, respectively. Therefore, we have

(nx - 1) (ny - 1) grouped points and (nx + ny - 1) ungrouped points, where nx = Nx - 1 and ny = Ny - 1.

Types of points in SEG for Nx = Ny = 8.

The estimation on this computational complexity is based on the arithmetic operations performed at each titeration for the additions/substractions (Add/Sub) and multiplications/divisions (Mul/Div) operations. Therefore, the number of operations required for SEG is given as in Table 1. The total number of arithmetic operations can be obtained by multiplying the number of arithmetic operations for each iteration with the number of iterations.

A quasi-newton algorithm

In this paper, the algorithm for large-scale nonlinear equations is designed by the following steps:

(i) A conjugate gradient (CG) algorithm is designed as a sub-algorithm to obtain the initial points of the main algorithm, where the sub-algorithm's initial point does not have any restrictions;

(ii) A quasi-Newton algorithm with the initial points given by sub-algorithm is defined as main algorithm, where a new nonmonotone line search technique is presented to get the step length αk. The given nonmonotone line search technique can avoid computing the Jacobian matrix. The global convergence and the 1 + q-order convergent rate of the main algorithm are established under suitable conditions. Numerical results show that the proposed method is competitive with a similar method for large-scale problems.

Consider the following nonlinear equations:

$$e(x) = 0, x \in Rn. \text{-}31.1$$

Here e:Rn → Rn is continuously differentiable and n denotes large-scale dimension. The large-scale nonlinear equations are difficult to solve since the relations of the variables x are complex and the dimension is larger. Problem (1.1) can model many real-life problems, such as engineering problems, dimensions of mechanical linkages, concentrations of chemical species, cross-sectional properties of structural elements, etc. If the Jacobian ∇e(x) of e is symmetric, then problem (1.1) is called a system of symmetric nonlinear equations. Let p be the norm function with p(x) = ½‖e(x)‖2, where ‖ · ‖ is the Euclidean norm. Then (1.1) is equivalent to the following global optimization models:

$$\min p(x), x \in Rn. \text{-}31.2$$

In fact, there are many actual problems that can convert to the above problems (1.2) (see [7-15] etc.) and have similar models (see [16-33] etc.). The iterative formula for (31.1) is

$$xk+1 = xk + \alpha kdk,$$

where αk is a step length and dk is a search direction. Now let us review some methods for αk and dk, respectively:

(i) Li and Fukashima [34] proposed an approximately monotone technique for αk:

$$p(xk + \alpha kdk)\text{-}p(xk) \leq -\delta 1\|\alpha kdk\|2 - \delta 2\|\alpha kek\|2 + \epsilon k\|ek\|2, \text{-}31.3$$

where ek = e(xk), δ1 > 0, δ2 > 0 are positive constants, αk = rik, r ∈ (0, 1), ik is the smallest nonnegative integer i satisfying (1.3) and ϵk such that

$$\sum k=0\infty \epsilon k<\infty. \text{-}31.4$$

(ii) Gu et al. [35] presented a descent line search technique:

$$p(xk + \alpha kdk)\text{-}p(xk) \leq -\delta 1\|\alpha kdk\|2 - \delta 2\|\alpha kek\|2. \text{-}31.5$$

(iii) Brown and Saad [36] gave the following technique to obtain αk:

$$p(xk + \alpha kdk)\text{-}p(xk) \leq \beta \alpha k\nabla p(xk)Tdk, \text{-}31.6$$

where β ∈ (0, 1) and ∇p(xk) = ∇e(xk)e(xk).

(iv) Based on this technique, Zhu [37] proposed a nonmonotone technique:

$$p(xk + \alpha kdk)\text{-}p(xl(k)) \leq \beta \alpha k\nabla p(xk)Tdk, \text{-}31.7$$

$$p(xl(k)) = \max 0\leq j\leq m(k)\{p(xk\text{-}j)\},\ m(0) = 0, \text{ and } m(k) = \min\{m(k - 1) + 1, M\},\ k \geq 1,$$

and M is a nonnegative integer.

(v) Yuan & Lu [38] gave a new technique:

$$p(xk+\alpha kdk)\text{-}p(xk)\leq \beta \alpha k2e(xk)Tdk, \text{-}31.8$$

and some convergence results are obtained.

Next we present some techniques for the calculation of dk. At present, there exist many well-known methods for dk, such as the

Newton method, the trust region method, and the quasi-Newton method, etc.

(i) The Newton method has the following form to get dk:

$$\nabla e(xk)dk = -e(xk).-31.9$$

This method is regarded as one of the most effective methods. However, its efficiency largely depends on the possibility to efficiently solve (1.9) at each iteration. Moreover, the exact solution of the system (1.9) could be too burdensome when the iterative point xk is far from the exact solution [39]. In order to overcome this drawback, inexact quasi-Newton methods are often used.

(ii) The quasi-Newton method is of the form

$$Bkdk + ek = 0, -31.10$$

where Bk is generated by a quasi-Newton update formula, where the BFGS (Broyden-Fletcher-Goldfarb-Shanno) update formula is one of the well-known quasi-Newton formulas with

Bk+1=Bk-BkskskTBkskTBksk+ykykTykTsk,-31.11

where sk = xk+1 - xk, yk = ek+1 - ek, and ek+1 = e(xk+1). Set Hk to be the inverse of Bk, then the inverse formula of (1.11) has the following form:

$$Hk+1=Hk-ykT(sk-Hkyk)skskT(ykTsk)2+(sk-Hkyk)skT+sk(sk-Hkyk)T(ykTsk)2=(I-skykTykTsk)Hk(I-yksk TykTsk)+skskTykTsk.-31.12$$

There exist many quasi-Newton methods see [37,38,40-45] representing the basic approach underlying most of the Newton-type large-scale algorithms.

The earliest nonmonotone line search framework was developed by Grippo et al. in [46] for Newton's methods. Many subsequent papers have exploited nonmonotone line search techniques of this nature (see [47-50] etc.), which shows that the nonmonotone technique works well in many cases. Considering these points, Zhu [37] proposed the nonmonotone line search (1.7). From (1.7), we can see that the Jacobian matrix $\nabla e(x)$ must be computed at every titeration. Computing the Jacobian matrix $\nabla e(x)$ may be expensive if n is large and for any n at every iteration. Thus, one might prefer to remove the matrix, leading to a new nonmonotone technique.

Inspired by the above observations, we make a study of inexact quasi-Newton methods with a new nonmonotone technique for solving smooth nonlinear equations. In the kth iteration of our algorithm, the following new nonmonotone technique is used to obtain αk:

$$p(xk + \alpha kdk) \leq p(xl(k)) + \alpha k\sigma e(xk)Tdk, -31.13$$

where $\sigma \in (0, 1)$ is a constant and dk is a solution of (1.10). Comparing with (1.7), the new technique (1.13) does not compute the Jacobian matrix $\nabla e(x)$. Then the storage and workload can be saved in theory. In Section 3, we will state the technique (1.13) is well defined.

It is well known that the initial point plays an important role in an algorithm. For example, the local superlinear convergence needs the iteration point x lies in the neighborhood of the optimal solution $x*$, if the choice of the point x is correct then the Newton method can get the optimal solution $x*$ just need one step, moreover, the correct initial point can speed up the efficiency of an algorithm. The nonlinear conjugate gradient method is one of the most effective line search methods for unconstrained optimization problems due to its simplicity and low memory requirement, especially for large-scale problems. Many scholars have made many studies and obtained lots of achievements on the CG methods or other similar new methods (see [51-55] etc.), where the results of [52] are especially interesting. It has been proved that the numerical performance of the CG methods is very interesting for large-scale problems in different application fields. These considerations prompt us to design a CG algorithm (sub-algorithm) for solving large-scale nonlinear equations, where the terminated iteration point of the CG algorithm was used as the initial point of the given algorithm (main algorithm). Then there exist two advantages from this process: one is that we can use the CG's characteristic to get a better initial point and another is that the good convergent results of the main algorithm can be preserved. The main attributes of this paper are stated as follows:

A sub-algorithm is designed to get the initial point of the main algorithm.

A new nonmonotone line search technique is presented; moreover, the Jacobian matrix ∇ek must not be computed at every titeration.

The given method possesses the sufficient descent property for the normal function p(x).

The global convergence and the 1 + q-order convergent rate of the new method are established under suitable conditions.

Numerical results show that this method is more effective than other similar methods.

We organize the paper as follows. In Section 2, the algorithms are stated. Convergent results are established in Section 3 and numerical results are reported in Section 4. In the last section, our conclusion is given. Throughout this paper, we use these notations: $\|\cdot\|$ is the Euclidean norm, e(xk) and g(xk+1) are replaced by ek and gk+1, respectively.

A Poisson-Nernst-Planck Equations

$$\frac{\partial c_\pm}{\partial t} = D_\pm \nabla \cdot \left[\nabla c_\pm + \frac{1}{k_B T} [\pm e(c_\pm \nabla \phi) + (c_\pm \nabla U)] \right] \text{in} \Omega_s,$$

$$\frac{\partial c_\pm}{\partial t} = D_\pm \nabla \cdot \left[\nabla c_\pm + \frac{1}{k_B T} c_\pm \nabla V_\pm + a^3 \left(\frac{c_\pm \nabla (c_+ + c_-)}{1 - c_+ a^3 - c_- a^3} \right) \right] \text{in} \Omega_s,$$

$$\frac{\partial c_\pm}{\partial t} = D_\pm \nabla \cdot \left[\nabla c_\pm + \frac{1}{k_B T} c_\pm \nabla V_\pm + N_\alpha(c_\pm) \right] \text{in} \Omega_s,$$

$$N_\alpha(c_\pm) = \alpha \left(a^3 \frac{c_\pm \nabla (c_+ + c_-)}{1 - c_+ a^3 - c_- a^3} \right), \alpha = \begin{cases} 0, & \text{forPNPE,} \\ 1, & \text{forMPNPE.} \end{cases}$$

$$J = \sum_\pm \pm \int_\Gamma e D_\pm n \cdot \left[\nabla c_\pm + \frac{1}{k_B T} [\pm e(c_\pm \nabla \phi) + c_\pm \nabla V] \right] ds.$$

$$\left[\nabla c_\pm + \frac{1}{k_B T} c_\pm \nabla V_\pm + N_\alpha(c_\pm) \right] \cdot n = 0 \text{on} \partial \Omega_{s,n},$$

$$\frac{\partial}{\partial t} \int_{\Omega_s} c_\pm dx = \int_{\Omega_s} D_\pm \nabla \cdot \left[\nabla c_\pm + \frac{1}{k_B T} c_\pm \nabla V_\pm + N_\alpha(c_\pm) \right] dx = \int_{\partial \Omega_{s,n}} \left[\nabla c_\pm + \frac{1}{k_B T} c_\pm \nabla V_\pm + N_\alpha(c_\pm) \right] \cdot n ds = 0.$$

$$\int_{\Omega_s} \left(\frac{\partial c_\pm}{\partial t} - D_\pm \nabla \cdot \left[\nabla c_\pm + \frac{1}{k_B T} c_\pm \nabla V_\pm + N_\alpha(c_\pm) \right] \right) \cdot v_\pm dx = 0,$$

$$\Rightarrow \int_{\Omega_s} \left(\frac{\partial c_\pm}{\partial t} \cdot v_\pm + D_\pm \left[\nabla c_\pm \frac{1}{k_B T} c_\pm \nabla V_\pm + N_\alpha(c_\pm) \right] \cdot \nabla v_\pm \right) dx = 0$$

$$\int_{\Omega_s} \left((c_\pm^{n+1} - c_\pm^n) \cdot v_\pm + \Delta t D_\pm \left[\nabla c_\pm^{n+1} + \frac{1}{k_B T} c_\pm^{n+1} \nabla V_\pm + N_\alpha(c_\pm^{n+1}) \right] \cdot \nabla v_\pm \right) dx = 0$$

$$c_\pm^k$$

$$\langle F_\alpha(c_\pm^{n+1}), v_\pm \rangle$$

$$\langle F_\alpha(c_\pm^{n+1}), v_\pm \rangle = \int_{\Omega_s} \left((c_\pm^{n+1} - c_\pm^n) \cdot v_\pm + \Delta t D_\pm \left[\nabla c_\pm^{n+1} + \frac{1}{k_B T} c_\pm^{n+1} \nabla V_\pm + N_\alpha(c_\pm^{n+1}) \right] \cdot \nabla v_\pm \right) dx.$$

$$\langle F_\alpha(c_\pm^{n+1}), v_\pm \rangle$$

$$c_\pm^{n+1}$$

$$\langle DF_\alpha(c_\pm^{n+1}) w_\pm, v_\pm \rangle = \frac{d}{d\tau} \langle F_\alpha(c_\pm^{n+1} + \tau w_\pm), v_\pm \rangle |_{\tau=0} = \int_{\Omega_s} \left(w_\pm \cdot v_\pm + \Delta t D_\pm \left[\nabla w_\pm + \frac{1}{k_B T} w_\pm \nabla V_\pm \right] \cdot \nabla v_\pm + DN_\alpha(c_\pm^{n+1}) \right) dx,$$

$$DN_\alpha(c_\pm^{n+1}) = \alpha \frac{a^3}{\left(1 - c_+^{n+1} a^3 - c_-^{n+1} a^3\right)^2} \left[a^3 (w_+ + w_-) c_\pm^{n+1} \nabla (c_+^{n+1} + c_-^{n+1}) - (a^3 (c_+^{n+1} + c_-^{n+1}) - 1) (\nabla (c_+^{n+1} + c_-^{n+1}) w_\pm + \nabla (w_+ + w_-) c_\pm^{n+1}) \right] \cdot \nabla v_\pm.$$

$$\langle DF_\alpha(c_\pm^{n+1}) w_\pm, v_\pm \rangle = - \langle F_\alpha(c_\pm^{n+1}), v_\pm \rangle$$

$$c_\pm^{n+1}$$

$$c_\pm^{n+1} = c_\pm^{n+1} + w_\pm.$$

$$\langle DF_\alpha(c_\pm^{n+1}) w_\pm, v_\pm \rangle$$

$$\langle F_\alpha(c_\pm^{n+1}) v_\pm \rangle$$

$$c_\pm^{n+1}$$

$$L_\alpha(c_\pm) = D_\pm \nabla \cdot \left[\nabla c_\pm + \frac{1}{k_B T} c_\pm \nabla V_\pm + N_\alpha(c_\pm) \right],$$

$$b_\pm = - \frac{D_\pm}{k_B T} \nabla V_\pm.$$

$$\sigma_\pm = \frac{h_\tau}{2 \| b_\pm \|_2} \psi(\text{Pe}_\tau),$$

$$\text{Pe}_\tau = \left(\frac{0.33 \| b_\pm \|_2 h_\tau}{2 D_\pm} \right) \text{and} \psi(q) = \begin{cases} 1, & \text{if} q > 1, \\ q, & \text{otherwise.} \end{cases}$$

$$\int_{\Omega_s} \left(\frac{\partial c_\pm}{\partial t} - L_\alpha(c_\pm) \right) \cdot v_\pm dx = 0.$$

$$\langle F_{\alpha,supg}(c_\pm), v_\pm \rangle = \int_{\Omega_s} \underbrace{\left(\frac{\partial c_\pm}{\partial t} - L_\alpha(c_\pm) \right) \cdot v_\pm dx}_{\text{weak form}} + \sum_{\tau \in T} \underbrace{\int_\tau \left(\frac{\partial c_\pm}{\partial t} - L_\alpha(c_\pm) \right) \cdot v_{\pm supg} dx}_{\text{stabilization}} = 0,$$

$$c_{\pm}^{n+1} = c_{\pm}^{n+1} + w_{\pm}$$

$$\phi \in \phi_d + H_0^1(\Omega)$$

$$v \in H_0^1(\Omega)$$

$$\rho^f = \sum_{i=i}^{N} q_i \delta(x - x_i).$$

$$H_0^1(\Omega)$$

$$\phi(x) = \begin{cases} \phi^r(x) + \phi^s(x) + \phi^h(x), & x \in \Omega_{mol}, \\ \phi^r(x), & x \in \Omega_s \cup \Omega_m. \end{cases}$$

$$\left[\left[\epsilon \frac{\partial \phi^r}{\partial n}\right]\right] = \epsilon_{mol} \frac{\partial\left(\phi^s + \phi^h\right)}{\partial n}, \text{ on } \partial\Omega_{mol}.$$

$$[[f(x)]] = \lim_{a \to 0} f(x + an) - f(x - an).$$

$$\langle g\left(\phi^h\right), v \rangle = \int_{\partial\Omega_{mol}} \epsilon_{mol} \frac{\partial\left(\phi^s + \phi^h\right)}{\partial n} v dx.$$

$$\frac{1}{N}\sum_{i=1}^{N} \left|\left(c_{\pm}^{n+1} - c_{\pm}^{n}\right) / \Delta t\right| \leq \theta,$$

A Stabilized Finite Element Method for Modified Poisson-Nernst-Planck Equations to Determine Ion Flow through a Nanopore.

The conventional Poisson-Nernst-Planck equations do not account for the finite size of ions explicitly. This leads to solutions featuring unrealistically high ionic concentrations in the regions subject to external potentials, in particular, near highly charged surfaces. A modified form of the Poisson-Nernst-Planck equations accounts for steric effects and results in solutions with finite ion concentrations. Here, we evaluate numerical methods for solving the modified Poisson-Nernst-Planck equations by modeling electric field-driven transport of ions through a nanopore. We describe a novel, robust finite element solver that combines the applications of the Newton's method to the nonlinear Galerkin form of the equations, augmented with stabilization terms to appropriately handle the drift-diffusion processes. To make direct comparison with particle-based simulations possible, our method is specifically designed to produce solutions under periodic boundary conditions and to conserve the number of ions in the solution domain. We test our finite element solver on a set of challenging numerical experiments that include calculations of the ion distribution in a volume confined between two charged plates, calculations of the ionic current though a nanopore subject to an external electric field, and modeling the effect of a DNA molecule on the ion concentration and nanopore current.

Governing equations

We consider the Poisson-Nernst-Planck Equations (PNPE) for a 1:1 electrolyte solution (referred to as solvent) described over a computational domain, denoted by Ω=Ωs∪Ωm, which includes both the solvent region, represented as Ωs, as well as a molecular or membrane region, Ωm, which is void of solvent. The time dependent PNPE are given as [33]

∂c±∂t=D±∇·[∇c±+1kBT[±e(c±∇ϕ)+(c±∇U)]]inΩs,-(42.1)

-∇·ϵ∇ϕ=e(c+-c-)inΩ,-(42.2)

where ϕ is the electrostatic potential and U is the potential due to other interactions (such as van der Waals and solvation forces), which is assumed to be the same for both ionic species. Hereafter, we will refer to potential U as a non-electrostatic potential, to differentiate it from the explicit electrostatic potential ϕ. In the Nernst-Planck equation, (2.1), the concentration of positive and negative ions are c+ and c-, respectively, kB is the Boltzmann's constant, T is temperature, e is the charge on an electron and D± are the diffusivities of the positive and negative ions, respectively. In the Poisson equation, (2.2), we assume a piecewise constant dielectric coefficient ε that is defined in the two sub-domains, Ωs and Ωm. For simplicity, we write the total potential energy experienced by an ion as V± =±eϕ+U.

The modified form of the PNPE (MPNPE) adds a nonlinear term to each of the two Nernst-Planck equations in (2.1) to model the steric repulsion. The Poisson equation remains un-changed, however the modified Nernst Planck equations are [35],

∂c±∂t=D±∇·[∇c±+1kBTc±∇V±+a3(c±∇(c++c-)1-c+a3-c-a3)]inΩs,-(42.3)

Here is the size of the ion (assumed to be the same for both species). As a result, in this model the maximum permitted concentration is bounded by 1/a3, which we refer to as the steric limit. To simplify the presentation of the material that follows, we write the PNPE

and the MPNPE as

$$\partial c\pm\partial t=D\pm\nabla\cdot[\nabla c\pm+1kBTc\pm\nabla V\pm+N\alpha(c\pm)]in\Omega s,-(42.4)$$

where

$$N\alpha(c\pm)=\alpha(a3c\pm\nabla(c++c-)1-c+a3-c-a3),\alpha=\{0,forPNPE,1,forMPNPE.\}-(42.5)$$

Description of the model system

A primary focus of this paper is the application of MPNPE solver to nanopores, wherein we compute the ionic current through a pore in a solid-state membrane. The domain we consider is depicted. Here, solution reservoirs above and below the membrane are connected through a nanopore, allowing positive and negative ions to pass from one side of the membrane to the other. We also consider a system where a DNA molecule is present inside the pore. Thus, the membrane (and the DNA, if present) comprises the domain Ωm, whereas the ionic solution, which consists of the solution reservoirs above and below the membranes and the nanopore, comprise the domain Ωs.

With the concentration profiles of (2.4), one important quantity is the ionic current J through a surface G with normal n, which is defined as

$$J=\sum\pm\pm\int\Gamma eD\pm n\cdot[\nabla c\pm+1kBT[\pm e(c\pm\nabla\phi)+c\pm\nabla V]]ds.-(42.6)$$

For example, in the 2D cross-section of the problem domain shown, we measure the ionic current through the plane in the middle of the pore, denoted by a dotted line.

The Poisson portion of the PNPE in (2.2) is solved with Dirichlet boundary conditions specified by ϕt and ϕb at the top and the bottom of the domain, and periodic boundary conditions along the other four sides. Further, the (unmodified and modified) Nernst-Planck equations in (2.1) and (2.3) use blocking boundary conditions on the interface of the membrane and the ionic solution, which is denoted ∂Ωs,n and is displayed with a dotted line while periodic boundary conditions are set at the remaining boundaries. Specifically, we consider blocking boundary conditions of the form

$$[\nabla c\pm+1kBTc\pm\nabla V\pm+N\alpha(c\pm)]\cdot n=0on\partial\Omega s,n,-(42.7)$$

where n is the unit normal on the surface ∂Ωs,n. One consequence of the blocking boundary conditions is that the integral of the concentration remains constant. That is, the total number of ions of each ion species is conserved:

$$\partial\partial t\int\Omega sc\pm dx=\int\Omega sD\pm\nabla\cdot[\nabla c\pm+1kBTc\pm\nabla V\pm+N\alpha(c\pm)]dx=\int\partial\Omega s,n[\nabla c\pm+1kBTc\pm\nabla V\pm+N\alpha(c\pm)]\cdot nds=0-(42.83)$$

Numerical methods

In this section we describe a numerical approach to solving the MPNPE. For the modified Nernst-Planck equations, we use the finite element method for spatial discretization, and backward Euler for time discretization due to the stiff nature of the solutions. Similarly, we also use a finite element discretization for solving the Poisson's equation. Our approach is used to solve for both the transient and steady state solutions. If the goal of the simulation is the steady state, then we evolve the solution until the measured temporal change reaches a desired tolerance. We have slightly different schemes for the transient and steady state solutions. For the transient solution, where accuracy at each time step is important, we solve the (2.3) and the (2.2) in a self-consistent manner, employing Gummel iteration. That is, at each time step, given some initial concentrations c± and electrostatic potential ϕ, we solve (2.3) using the finite element method. The updated values of the concentrations are then used to solve (2.2), and this process is repeated until convergence. Then we move on to the next time step. This is illustrated in Equation 4. Overview of the numerical scheme. At each time step, we iterate between the solves for the modified Nernst-Planck equations and the Poisson equation.

For the steady state method, at each time step, we solve the (2.3), with the updated values of the concentration used in the solution of (2.2). However, further iteration is not needed and the updated value of the electric potential is immediately used as input to (2.3) for the next time step. The steady state method corresponds to employing one iteration in the Gummel iteration - i.e., we do not iterate to convergence at each time step. This process works well in practice since the primary purpose of the steady state solution are the final concentrations and is similar to previous strategies in the setting of the PNPE and MPNPE [34,41,42]. However, our scheme has two distinct features: 1) it uses a stabilized method, and 2) it accounts for the conservation of ions. We now proceed to describe our method in more detail (the approach is implemented in the finite element package Dolfin [43]. In the following we devise two finite element methods: one for the case in which there is only an electric field, and one in which a non-electric potential is also applied. We seek to highlight the benefits of incorporating the steric effects into the modified Nernst-Planck equations. Hence we derive our finite element method for both PNPE and MPNPE, highlighting several important differences. Themethodswe propose are used further when we devise a method for solving the MPNPE in the presence of an applied non-electrostatic potential.

7.3.1.Non-linear finite element method for modified Nernst-Planck equations: For the (modified) Nernst-Planck equations, we seek a solution in H1(Ωs) × H1(Ωs) for c±, where H1 is a Sobolev space with standard notation [207]. The weak form of (2.4) (which represents both (2.1) and (2.3) is found by integrating against a test function v± = (v+,v-) ∈ [H1(Ωs)]2 : find c± ∈ [H1(Ωs)]2 such that

$$\int\Omega s(\partial c\pm\partial t-D\pm\nabla\cdot[\nabla c\pm+1kBTc\pm\nabla V\pm+N\alpha(c\pm)])\cdot v\pm dx=0,-(43.1a)$$

$$\Rightarrow\int\Omega s(\partial c\pm\partial t\cdot v\pm+D\pm[\nabla c\pm1kBTc\pm\nabla V\pm+N\alpha(c\pm)]\cdot\nabla v\pm)dx=0\text{-}(43.1b)$$

for all v±in[H1(Ωs)]2. Note that we used integration by parts along with the blocking and periodic boundary conditions to derive (3.1b) from (3.1a). From this, we apply the backward Euler method in time to (3.1b) and multiply by Dt to arrive at

$$\int\Omega s((c\pm n+1\text{-}c\pm n)\cdot v\pm+\Delta tD\pm[\nabla c\pm n+1+1kBTc\pm n+1\nabla V\pm+N\alpha(c\pm n+1)]\cdot\nabla v\pm)dx=0\text{-}(43.2)$$

for all v± ∈ [H1(Ωs)]2. Here, c±k are the concentrations at time step k, while Δt is the length of the time step.

For the PNPE, (3.2) is a pair of uncoupled equations, linear in the unknown variables c±. For the MPNPE, (3.2) is a pair of coupled non-linear equations, and hence result in a more complicated form than PNPE. To address the nonlinearity, we use a straightforward application of Newton's method to find the concentrations c± for the MPNPE. Specifically, we define the form in (3.2) as

⟨Fα(c±n+1),v±⟩:

$$\langle F\alpha(c\pm n+1),v\pm\rangle=\int\Omega s((c\pm n+1\text{-}c\pm n)\cdot v\pm+\Delta tD\pm[\nabla c\pm n+1+1kBTc\pm n+1\nabla V\pm+N\alpha(c\pm n+1)]\cdot\nabla v\pm)dx.\text{-}(43.3)$$

Here we again use the α notation to write the weak form for both the NP and MNP equations.

The MNPE in (3.3) are nonlinear in c± and linear in v±. The first step in a Newton's method is to linearize this form with respect to c±, to arrive at a bilinear form in w± and v±. This is accomplished by taking a variational derivative of ⟨Fα(c±n+1),v±⟩ with respect to c±n+1,

$$\langle DF\alpha(c\pm n+1)w\pm,v\pm\rangle=dd\tau\langle F\alpha(c\pm n+1+\tau w\pm),v\pm\rangle|\tau=0,=\int\Omega s(w\pm\cdot v\pm+\Delta tD\pm[\nabla w\pm+1kBTw\pm\nabla V\pm]\cdot\nabla v\pm+DN\alpha(c\pm n\ 1))dx,\text{-}(43.4)$$

where

$$DN\alpha(c\pm n+1)=\alpha a3(1\text{-}c+n+1a3\text{-}c\text{-}n+1a3)2[a3(w++w\text{-})c\pm n+1\nabla(c+n+1+c\text{-}n+1)\text{-}(a3(c+n+1+c\text{-}n+1)\text{-}1)(\nabla(c+n+1+c\text{-}n+1)w\pm+\nabla(w++w\text{-})c\pm n+1)]\cdot\nabla v\pm.\text{-}(43.5)$$

As a consequence, with α=0 for PNPE, the derivative DN0 is absent from (3.4).

With the derivative defined, we arrive at Newton's method, which first computes w± by solving

$$\langle DF\alpha(c\pm n+1)w\pm,v\pm\rangle=\text{-}\langle F\alpha(c\pm n+1),v\pm\rangle\text{-}(3.6)$$

followed by the estimate of c±n+1 with

$$c\pm n+1=c\pm n+1+w\pm.\text{-}(43.7)$$

Solving the weak problem in (3.6) reduces to a linear system of the form Aw = f, where stiffness matrix A is formed by ⟨DFα(c±n+1) w±,v±⟩, the force vector f by ⟨Fα(c±n+1)v±⟩, and where w represents the unknown coefficients for the function w±.

In the case of the NP equations, with α = 0, the form DFα in (3.6) does not depend on c±n+1, and the Newton's method converges in one iteration (as we expect for linear problems). For simplicity, we drop the arguments to the forms DFα and Fα. Then DF0 and DF1 represent the forms in (3.4). Similarly F0 and F1 are the forms for NP and MMP in (3.3). The relationship between these forms is expressed as,

$$DF1=DF0+DN1,\text{-}(43.8)$$

$$F1=F0+N1\cdot\nabla v\pm,\text{-}(43.9)$$

which allows us to view the Newton step for MNP as,

$$DF0+DN1=\text{-}(F0+N1\cdot\nabla v\pm).\text{-}(43.10)$$

From this equation, we notice that the Newton step for the MPNPE is exactly the Newton step for the PNPE with the addition of stabilizing terms DN1(c±) and N1(c±)•∇v±.

$$\sum_{k=0}^{\infty} c_k < \infty.$$

$$p(x_k + \alpha_k d_k) - p(x_k) \leq \beta\alpha_k^2 e(x_k)^T d_k.$$

$$B_{k+1} = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k y_k^T}{y_k^T s_k}.$$

$$H_{k+1} = H_k - \frac{y_k^T (s_k - H_k y_k) s_k s_k^T}{(y_k^T s_k)^2} + \frac{(s_k - H_k y_k) s_k^T + s_k (s_k - H_k y_k)^T}{(y_k^T s_k)^2}$$

$$= \left(I - \frac{s_k y_k^T}{y_k^T s_k}\right) H_k \left(I - \frac{y_k s_k^T}{y_k^T s_k}\right) + \frac{s_k s_k^T}{y_k^T s_k}.$$

$$e_k^T d_k < 0$$

$$\nabla p(x_k)^T d_k = e(x_k)^T \nabla e(x_k) d_k$$

$$
\begin{aligned}
&= e(x_k)^T[(\nabla e(x_k) - B_k)d_k - e(x_k)] \\
&= e(x_k)^T(\nabla e(x_k) - B_k)d_k - e(x_k)^T e(x_k). \\
\nabla p(x_k)^T d_k + \|e(x_k)\|^2 &= e(x_k)^T(\nabla e(x_k) - B_k)d_k \\
&\le \|e(x_k)\|\|(\nabla e(x_k) - B_k)d_k\|. \\
\nabla p(x_k)^T d_k &\le \|e(x_k)\|\|(\nabla e(x_k) - B_k)d_k\| - \|e(x_k)\|^2 \\
&\le -(1 - \epsilon_w)\|e(x_k)\|^2.
\end{aligned}
$$

$$\lim_{k\to\infty} \| e_k \| = 0.$$

$$p(x_{k+1}) - p(x_{l(k)}) \le \sigma\alpha_k e_k^T d_k < 0.$$

$$b_2\|d_k\|^2 \le d_k^T B_k d_k = -e_k^T d_k \le b_1\|d_k\|^2.$$

$$
\begin{aligned}
p(x_{l(k)}) &= p(x_{l(k)-1} + \alpha_{l(k)-1}d_{l(k)-1}) \\
&\le \max_{0\le j\le m(l(k)-1)}\{p(x_{l(k)-j-1})\} + \sigma\alpha_{l(k)-1}g_{l(k)-1}^T d_{l(k)-1} \\
&\le \max_{0\le j\le m(l(k)-1)}\{p(x_{l(k)-j-1})\} - \sigma b_2\alpha_{l(k)-1}\|d_{l(k)-1}\|^2.
\end{aligned}
$$

$$\lim_{k\to\infty}\alpha_{l(k)-1}\|d_{l(k)-1}\|^2 = 0.$$

$$\lim_{k\to\infty} \inf d_{l(k)-1} = 0$$

$$\lim_{k\to\infty} \inf \alpha_{l(k)-1} = 0.$$

$$\lim_{k\to\infty} \| d_{l(k)-j} \| = 0$$

$$\lim_{k\to\infty} p(x_{l(k)-j}) = \lim_{k\to\infty} p(x_{l(k)})$$

$$\lim_{k\to\infty} p(x_{l(k)}) = \lim_{k\to\infty} p(x_k).$$

$$p(x_{k+1}) - p(x_{l(k)}) \le \alpha_k\sigma e_k^T d_k \le \alpha_k\sigma b_2\|d_k\|^2.$$

$$\lim_{k\to\infty}\alpha_k\|d_k\|^2 = 0,$$

$$\lim_{k\to\infty}\alpha_k = 0$$

$$\lim_{k\to\infty} \| d_k \| = 0.$$

$$\alpha_k' = \frac{\alpha_k}{r}$$

$$p(x_k + \alpha_k' d_k) - p(x_k) \ge p(x_k + \alpha_k' d_k) - p(x_{l(k)}) > \sigma\alpha_k' e_k^T d_k.$$

$$p(x_k + \alpha_k' d_k) - p(x_k) = \alpha_k' \nabla p(x_k)^T d_k + o(\alpha_k' \| d_k \|).$$

$$\nabla p(x_k)^T d_k = e_k^T \nabla e(x_k) d_k \le \delta'' e_k^T d_k.$$

$$[\delta'' - \sigma]\alpha_k' e_k^T d_k + o(\alpha_k' \| d_k \|) \ge 0.$$

$$e_k^T d_k < 0$$

$$\alpha_k' \| d_k \|$$

$$\lim_{k\to\infty} \frac{e_k^T d_k}{\|d_k\|} = 0.$$

$$\lim_{k\to\infty} \| d_k \| = 0.$$

$$a \equiv \{\| \nabla e(x^*)\| + \tfrac{1}{2c}, 2c\}, c = \| \nabla e(x^*)^{-1} \|.$$

$$\tfrac{1}{a}\|x_k - x^*\| \le \|e(x_k)\| \le a\|x_k - x^*\|$$

$$
\begin{aligned}
&B_k(x_{k+1} - x^*) + [\nabla e(x_k)(x_k - x^*) - B_k(x_k - x^*)] \\
&+ [e(x_k) - e(x^*) - \nabla e(x_k)(x_k - x^*)] \\
&= e(x_k) + B_k d_k = 0,
\end{aligned}
$$

$$
\begin{aligned}
f_1(x) &= e^{x_1} - 1, \\
f_i(x) &= \tfrac{i}{10}(e^{x_i} + x_{i-1} - 1), i = 2, 3, \ldots, n.
\end{aligned}
$$

$$x_0 = \left(\tfrac{1}{n^2}, \tfrac{1}{n^2}, \ldots, \tfrac{1}{n^2}\right)^T$$

$$f_i(x) = 2(n + i(1 - \cos x_i) - \sin x_i - \textstyle\sum_{j=1}^n \cos x_j)(2\sin x_i - \cos x_i), i = 1, 2, 3, \ldots, n.$$

$$x_0 = \left(\tfrac{101}{100n}, \tfrac{101}{100n}, \ldots, \tfrac{101}{100n}\right)^T$$

$$f_i(x) = \ln(x_i + 1) - \tfrac{x_i}{n}, i = 1, 2, 3, \ldots, n.$$

$$
\begin{aligned}
f_1(x) &= (3 - 0.5x_1)x_1 - 2x_2 + 1, \\
f_i(x) &= (3 - 0.5x_i)x_i - x_{i-1} + 2x_{i+1} + 1, \\
&\quad i = 2, 3, \ldots, n-1, \\
f_n(x) &= (3 - 0.5x_n)x_n - x_{n-1} + 1.
\end{aligned}
$$

$$
\begin{aligned}
f_1(x) &= 3x_1^3 + 2x_2 - 5 + \sin(x_1 - x_2)\sin(x_1 + x_2), \\
f_i(x) &= -x_{i-1}e^{x_{i-1} - x_i} + x_i(4 + 3x_i^2) + 2x_{i+1} \\
&\quad + \sin(x_i - x_{i+1})\sin(x_i + x_{i+1}) - 8, i = 2, 3, \ldots, n-1, \\
f_n(x) &= -x_{n-1}e^{x_{n-1} - x_n} + 4x_n - 3.
\end{aligned}
$$

$$h(x) = \textstyle\sum_{i=1}^n (e^{x_i} - x_i)$$

$$x_0 = \left(\tfrac{1}{n}, \tfrac{2}{n}, \ldots, 1\right)^T$$

$$h(x) = \textstyle\sum_{i=1}^n \tfrac{i}{10}(e^{x_i} - x_i)$$

$$
\begin{aligned}
f_i(x) &= \tfrac{i}{10}(e^{x_i} - 1), i = 1, 2, 3, \ldots, n. \\
f_i(x) &= x_i - 1, i = 1, 2, 3, \ldots, n-2, \\
f_{n-1}(x) &= \textstyle\sum_{j=1}^{n-2} j(x_j - 1), \\
f_n(x) &= \left(\textstyle\sum_{j=1}^{n-2} j(x_j - 1)\right)^2.
\end{aligned}
$$

$$x_0 = \left(1 - \tfrac{1}{n}, 1 - \tfrac{2}{n}, \ldots, 0\right)^T$$

$$f_1(x) = 2x_1 + 0.5h^2(x_1 + h)^3 - x_2,$$

$$f_i(x) = 2x_i + 0.5h^2(x_i + hi)^3 - x_{i-1} + x_{i+1}.$$
$$i = 2, 3, \ldots, n-1$$
$$f_n(x) = 2x_n + 0.5h^2(x_n + hn)^3 - x_{n-1}.$$
$$h = \frac{1}{n+1}.$$
$$e(x) = Ax + \frac{1}{(n+1)^2}F(x) = 0.$$
$$A = \begin{bmatrix} 8 & -1 & & & & \\ -1 & 8 & -1 & & & \\ & -1 & 8 & -1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & \ddots & \ddots & -1 \\ & & & & -1 & 8 \end{bmatrix}.$$
$$y_k^T s_k > 0$$
$$d_k^T e_k < 0$$
$$d_k^T \nabla e(x_k) e_k < 0$$
$$d_k = \begin{cases} -e_k + \frac{e_k^T(e_k - e_{k-1})}{\|e_{k-1}\|^2} d_{k-1} & \text{if } k \geq 1. \\ -e_k & \text{if } k = 0. \end{cases}$$
$$c_k = \frac{1}{Nl^2}$$

A stabilized finite element method for the modified Nernst Planck equations

The discretization scheme we have developed so far works well if there is no strong electrostatic potential (represented as ϕ) or applied non-electrostatic potential (i.e. U = 0). However, in the presence of a strong electorstatic potential or non-electrostatic applied potential, the Nernst-Planck equations have a large drift term, which is a challenge to standard Galerkin methods. Using the standard Galerkin approach results in a solution with spurious values [45] - e.g. the concentration becomes negative in portions of the domain, as illustrated earlier in Equation 1 for the nanopore system. We see that the concentration becomes negative in parts of the domain. One remedy is to augment the Galerkin weak form by adding artificial dissipative terms to stabilize the method. To this end, we use a variant of streamline upwind Petrov-Galerkin method (SUPG) for stabilizing our scheme in presence of steep gradients in ϕ or U [46,47]. We develop stabilized schemes for both the PNPE and the MPNPE. We develop two such schemes for the MPNPE. One arises from a standard application of SUPG to the MPNPE, whereas the other is developed by adding the nonlinear terms of the MPNPE to the SUPG scheme for the PNPE. The latter scheme, which we call “Fast SUPG” improves on the former, called the “Full SUPG” scheme, by increasing its computational efficiency, as we explain later. We use the “Fast SUPG” method in this paper, unless noted. The relationship between different SUPG schemes.

Relationship between SUPG methods

To simplify the presentation of our SUPG scheme, we introduce some notation. The differential operator is given as

$$L\alpha(c\pm)=D\pm\nabla\cdot[\nabla c\pm+1kBTc\pm\nabla V\pm+N\alpha(c\pm)],\text{-}(43.11)$$

which consists of a flow field governed by the applied potential V± and denoted by

$$b\pm=\text{-}D\pm kBT\nabla V\pm.\text{-}(43.12)$$

We also isolate the Péclet number Pe with a stability parameter of the form

$$\sigma\pm=h\tau 2\|b\pm\|2\psi(Pe\tau),\text{-}(43.13)$$

where hτ denotes the diameter of the element τ, and with

$$Pe\tau=(0.33\|b\pm\|2h\tau 2D\pm)and\psi(q)=\{1,ifq>1,q,otherwise.\}\text{-}(43.14)$$

The Péclet number of element τ, denoted as Peτ is an indication of the strength of advection. Specifically, a Péclet number greater than 1.0 indicates that advection is dominating the flow and that stabilization may be necessary, and we use the values developed in [45,48].

Using the above notation, we write (3.1a) as,

$$\int\Omega s(\partial c\pm\partial t\text{-}L\alpha(c\pm))\cdot v\pm dx=0.\text{-}(43.15)$$

From this, we define the SUPG weak form using integration-by-parts similar to the derivation in (3.1a), to arrive

$$\langle F\alpha,supg(c\pm),v\pm\rangle=\int\Omega s(\partial c\pm\partial t\text{-}L\alpha(c\pm))\cdot v\pm\square dxweak\ form+\sum\tau\in T\int\tau(\partial c\pm\partial t\text{-}L\alpha(c\pm))\cdot v\pm supg\square dx=0,stabilization\text{-}(43.16)$$

where the test functions in stabilized form are

v±supg=σ±b±·∇v±.-(43.17)

For the PNPE, (3.16) is solved as a pair of uncoupled equations. However, for the MPNPE, we use a nonlinear solver, as discussed in the previous section. The nonlinear SUPG scheme is similar to the approach taken in [49] for use with nonlinear Navier-Stokes equations. The convergence rate of the L2 error for the SUPG scheme is typically half an order less than the Galerkin method [50]. That is, the convergence rate of the L2 error is O (hk+1⁄2) for the SUPG scheme, and O (hk+1) for the Galerkin method, where k refers to the polynomial degree of the approximating space. However, the benefit of the SUPG method lies in its stability properties. The stability of the SUPG method yields a physically meaningful solution as compared to the Galerkin method, even for coarse meshes.

Current mathematical methods used in qsar_qspr studie

Recently, the mathematical methods applied to the regression of QASR/QSPR models are developing very fast, and new methods, such as Gene Expression Programming (GEP), Project Pursuit Regression (PPR) and Local Lazy Regression (LLR) have appeared on the QASR/QSPR stage. At the same time, the earlier methods, including Multiple Linear Regression (MLR), Partial Least Squares (PLS), Neural Networks (NN), support Vector Machine (SVM) and so on, were used to improve the performance in QASR/QSPR studies. These new and upgraded methods and algorithms are described in detail, and their advantages and disadvantages are evaluated and discussed, to show their application potential in QASR/QSPR studies in the future. As a common and successful research approach, quantitative structure activity/property relationship (QASR/QSPR) studies are applied extensively to chemometrics, pharmacodynamics, pharmacokinetics, toxicology and so on.

$$h_j(x) = \exp\left(\frac{-\|x-c_j\|^2}{r_j^2}\right)$$

$$E(y|x) \triangleq \hat{y}(x) = \frac{\int_{-\infty}^{+\infty} y f(x,y)dy}{\int_{-\infty}^{+\infty} f(x,y)dy}$$

$$f(x,y) = \frac{1}{n\sigma} \sum_{i=1}^{n} w\frac{(x-x_i)}{\sigma}$$

$$\hat{y}(x) = \frac{\sum_{i=1}^{n} y_i \exp(-D(x,x_i))}{\sum_{i=1}^{n} \exp(-D(x,x_i))}$$

$$D(x,x_i) = \sum_{j=1}^{p} \left(\frac{x_j - x_{ij}}{\sigma_j}\right)^2$$

$$Q_{LS-SVM} = \frac{1}{2} w^T w + \gamma \sum_{k=1}^{N} e_k^2$$

$$w = \sum_{k=1}^{N} \alpha_k x_k^T x + b \quad \text{with} \quad \alpha_k = 2\gamma e_k$$

$$y = \sum_{k=1}^{N} \alpha_k x_k^T x + b$$

$$\alpha_k = (x_k^T x + (2\gamma)^{-1}(y_k - b))$$

$$f(x) = \sum_{k=1}^{N} \alpha_k K(x,x_k) + b$$

$$\{01234567\}$$

Current Mathematical Methods Used in QSAR_QSPR Studie

Multiple linear regression (MLR)

MLR is one of the earliest methods used for constructing QSAR/QSPR models, but it is still one of the most commonly used ones to date. The advantage of MLR is its simple form and easily interpretable mathematical expression. Although utilized to great effect, MLR is vulnerable to descriptors which are correlated to one another, making it incapable of deciding which correlated sets may be more significant to the model. Some new methodologies based on MLR have been developed and reported in recent papers aimed at improving this technique. These methods include Best Multiple Linear Regression (BMLR), Heuristic Method (HM), Genetic Algorithm based Multiple Linear Regression (GA-MLR), Stepwise MLR, Factor Analysis MLR and so on. The three most important and commonly used of these methods are described in detail below.

Best multiple linear regressions (BMLR): BMLR implements the following strategy to search for the multi-parameter regression with the maximum predicting ability. All orthogonal pairs of descriptors i and j (with R2ij < R2min, default value R2ij < 0.1) are found in a given data set. The property analyzed is treated by using the two-parameter regression with the pairs of descriptors, obtained in the first step. The Nc (default value Nc = 400) pairs with highest regression correlation coefficients are chosen for performing the higher-order regression treatments. For each descriptor pair, obtained in the previous step, a non-collinear descriptor scale, k (with R2ik < R2nc and R2kj < R2nc, default value R2 < 0.6) is added, and the respective three-parameter regression treatment is performed. If the Fisher criterion at a given probability level, F, is smaller than that for the best two-parameter correlation, the latter is chosen as the final result. Otherwise, the Nc (default value Nc = 400) descriptor triples with highest regression correlation coefficients are chosen for the next step. For each descriptor set, chosen in the previous step, an additional non-collinear descriptor scale is added, and the respective (n + 1)-parameter regression treatment is performed. If the Fisher criterion at the given probability level, F, is smaller than for the best two-parameter correlation, the latter is chosen as the final result. Otherwise, the Nc (default value Nc = 400) sets descriptor sets with highest regression correlation coefficients are chosen, and this step repeated with n=n+1 [41]. As an improved method based on MLR, BMLR is instrumental for variable selection and QSAR/QSPR modeling [42-50]. Like MLR, BMLR is noted for its simple and interpretable mathematical expression. Moreover, overcoming the shortcomings of MLR, BMLR works well when the number of compounds in the training set doesn't exceed the number of molecular descriptors by at least a factor of five. However, BMLR will derive an unsatisfactory result when the structure-activity relationship is non-linear in nature. When too many descriptors are involved in a calculation, the modeling process will be time consuming. To speed up the calculations, it is advisable reject descriptors with insignificant variance within the dataset. This will significantly decrease the probability of including unrelated descriptors by chance. In addition, BMLR is unable to build a one-parameter model. BMLR is commercially available in the software packages CODESSA [49] or CODESSA PRO [50].

Heuristic method (HM): HM, an advanced algorithm based on MLR, is popular for building linear QSAR/QSPR equations because of its convenience and high calculation speed. The advantage of HM is totally based on its unique strategy of selecting variables. The details of selecting descriptors are as follows: first of all, all descriptors are checked to ensure that values of each descriptor are available for each structure. Descriptors for which values are not available for every structure in the data are discarded. Descriptors having a constant value for all structures in the data set are also discarded. Thereafter all possible one-parameter regression models are tested and the insignificant descriptors are removed. As a next step, the program calculates the pair correlation matrix of descriptors and further reduces the descriptor pool by eliminating highly correlated descriptors. The details of validating intercorrelation are: (a) all quasi-orthogonal pairs of structural descriptors are selected from the initial set. Two descriptors are considered orthogonal if their intercorrelation coefficient rij is lower than 0.1; (b) the pairs of orthogonal descriptors are used to compute the biparametric regression equations; (c) to a multi-linear regression (MLR) model containing n descriptors, a new descriptor is added to generate a model with n+1 descriptors if the new descriptor is not significantly correlated with the previous n descriptors; step (c) is repeated until MLR models with a prescribed number of descriptors are obtained. The goodness of the correlation is tested by the square of coefficient regression (R2), square of cross-validate coefficient regression (q2), the F-test (F), and the standard deviation (S) [41].

HM is commonly used in linear QSAR and QSPR studies, and also as an excellent tool for descriptor selection before a linear or nonlinear model is built [51,170-180]. The advantages of HM are the high speed and the absence of software restrictions on the size of the data set. HM can either quickly give a good estimation about what quality of correlation to expect from the data, or derive several best regression models. HM usually produces correlations 2-5 times faster than other methods with comparable quality. Additionally, the maximum number of parameters in the resulting model can be fixed in accordance with the situation so as to save time. As a method inherited from MLR, HM is also limited in linear models.

Genetic algorithm based multiple linear regressions (GA-MLR): Combining Genetic Algorithm (GA) with MLR, a new method called GA-MLR is becoming popular in currently reported QSAR and QSPR studies [181-195]. In this method, GA is performed to search the feature space and select the major descriptors relevant to the activities or properties of the compounds. This method can deal with z large search space efficiently and has less chance to become a local optimal solution than the other algorithms. We give a brief summary of the main procedure of GA herein. The first step of GA is to generate a set of solutions (chromosomes) randomly, which is called an initial population. Then, a fitness function is deduced from the gene composition of a chromosome [196-202]. The Friedman LOF function is commonly used as the fitness function, which was defined as follows:

$$LOF = \{SSE/(1-(c+dp/n))\}2-(51)$$

where SSE is the sum of squares of errors, c is the number of the basis function (other than the constant term), d is the smoothness

factor, p is the number of features in the model, and n is the number of data points from which the model is built. Unlike the R2 error, the LOF measure cannot always be reduced by adding more terms to the regression model. By limiting the tendency to simply add more terms, the LOF measure resists over-fitting of a model. Then, crossover and mutation operations are performed to generate new individuals. In the subsequent selection stage, the fittest individuals evolve to the next generation. These steps of evolution continue until the stopping conditions are satisfied. After that, the MLR is employed to correlate the descriptors selected by GA and the values of activities or properties. GA, a well-estimated method for parameter selection, is embedded in GA-MLR method so as to overcome the shortage of MLR in variable selection. Like the MLR method, the regression tool in GA-MLR, is a simple and classical regression method, which can provide explicit equations. The two parts have a complementation for each other to make GA-MLR a promising method in QSAR/QSPR research.

Partial least squares (PLS)

The basic concept of PLS regression was originally developed by Sammi Tand Li ZG [203,204]. As a popular and pragmatic methodology, PLS is used extensively in various fields. In the field of QSAR/QSPR, PLS is famous for its application to CoMFA and CoMSIA. Recently, PLS has evolved by combination with other mathematical methods to give better performance in QSAR/QSPR analyses. These evolved PLS', such as Genetic Partial Least Squares (G/PLS), Factor Analysis Partial Least Squares (FA-PLS) and Orthogonal Signal Correction Partial Least Squares (OSC-PLS), are briefly introduced in the following sections.

Genetic partial least squares (G/PLS): G/PLS is derived from two QSAR calculation methods Genetic Function Approximation (GFA) [205,206] and PLS. The G/PLS algorithm uses GFA to select appropriate basis functions to be used in a model of the data and PLS regression is used as the fitting technique to weigh the basis functions' relative contributions in the final model. Application of G/PLS thus allows the construction of larger QSAR equations while still avoiding over-fitting and eliminating most variables. As the regression method used in Molecular Field Analysis (MFA), a well-known 3D-QSAR analysis tool, G/PLS is commonly used. The recent literatures related to G/PLS are mainly listed as [207-217].

Factor analysis partial least squares (FA-PLS): This is the combination of Factor Analysis (FA) and PLS, where FA is used for initial selection of descriptors, after which PLS is performed. FA is a tool to find out the relationships among variables. It reduces variables into few latent factors from which important variables are selected for PLS regression. Most of the time, a leave-one-out method is used as a tool for selection of optimum number of components for PLS. We can find examples of FA-PLS used in QSAR analysis in [215, 71, 218-293].

Orthogonal signal correction partial least squares (OSC-PLS): Orthogonal signal correction (OSC) was introduced by Niazi [222] to remove systematic variation from the response matrix X that is unrelated, or orthogonal, to the property matrix Y. Therefore, one can be certain that important information regarding the analyte is retained. Since then, various OSC algorithms have been published in an attempt to reduce model complexity by removing orthogonal components from the signal. In abstracto, a preprocessing with OSC will help traditional PLS to obtain a more precise model, as proven in many studies of spectral analysis [223-234]. To date, unfortunately, there are only a few reports in which OSC-PLS is applied to QSAR/QSPR studies [89-91], but more QSAR or QSPR research involving application of the OSC-PLS method are expected in the future.

Neural networks (NN)

As an alternative to the fitting of data to an equation and reporting the coefficients derived therefrom, neural networks are designed to process input information and generate hidden models of the relationships. One advantage of neural networks is that they are naturally capable of modeling nonlinear systems. Disadvantages include a tendency to overfit the data, and a significant level of difficulty in ascertaining which descriptors are most significant in the resulting model. In the recent QSAR/QSPR studies, RBFNN and GRNN are the most frequently used ones among NN.

Radial basis function neural network (RBFNN): The RBFNN consists of three layers: an input layer, a hidden layer and an output layer. The input layer does not process the information; it only distributes the input vectors to the hidden layer. Each neuron on the hidden layer employs a radial basis function as a nonlinear transfer function to operate on the input data. In general, there are several radial basis functions (RBF): linear, cubic, thin plate spline, Gaussian, multi-quadratic and inverse multi-quadratic. The most often used RBF is a Gaussian function that is characterized by a center (cj) and width (rj). Due to the limited length of writing, only Gaussian RBF is introduced in this paper. The nonlinear transformation with RBF in the hidden layer is given as follows:

$$hj(x)=\exp(-\|x-cj\|2rj2)-(52)$$

where, hj is the notation for the output of the jth RBF unit, cj and rj are the center and width of the jth RBF, respectively. The operation of the output layer is linear, which is given as below:

$$yk(x)=\sum wkjhj(x)+bk-(53)$$

where, yk is the kth output unit for the input vector x, wkj is the weight connection between the kth output unit and the jth hidden layer unit, and bk is the bias. The training procedure when using RBF involves selecting centers, width and weights. In this paper, the forward subset selection routine was used to select the centers from training set samples. The adjustment of the connection weight between the hidden layer and the output layer was performed using a least squares solution after the selection of centers and widths of radial basis functions. Compared with the Back Propagation Neural Network (BPNN), RBFNN has the advantage of short training time

and is guaranteed to reach the global minimum of error surface during training process. The optimization of its topology and learning parameters are easy to implement. Applications of RBFNN in QSAR/QSPR studies can be found in [62,170,173,178,198,237-247].

General regression neural network (GRNN): GRNN, one of the so-called Bayesian networks, is a type of neural network using kernel-based approximation to perform regression. It was introduced by Specht [248]. GRNN is a nonparametric estimator that calculates a weighted average of the target values of training patterns by the probability density function using Parzen's nonparametric estimator. For GRNN, the predicted value is the most probable value E(y|x):

$$E(y|x)=\square y\hat{}(x)=\int_{-\infty}^{+\infty} yf(x,y)dy\int_{-\infty}^{+\infty} f(x,y)dy\text{-}(54)$$

where f(x, y) is the probability density function. This can be estimated from the training set by using the Parzen's nonparametric estimator [193]:

$$f(x,y)=1n\sigma\sum i=1nw(x\text{-}xi)\sigma\text{-}(55)$$

where n is the sample size, s is a scaling parameter that defines the width of the bell curve that surrounds each sample point, W(d) is a weighting function that has its largest value at d=0, and (x-xi) is the distance between the unknown sample and a data point. The Gaussian function is frequently used as the weighting function because it is well behaved, easily calculated, and satisfies the conditions required by Parzen's estimator. Substituting Parzen's nonparametric estimator for f(x, y) and performing the integrations leads to the fundamental equation of GRNN:

$$y\hat{}(x)=\sum i=1nyi\ \exp(-D(x,xi))\sum i=1nexp(-D(x,xi))\text{-}(56)$$

where:

$$D(x,xi)=\sum j=1p(xj\text{-}xij\sigma j)2\text{-}(57)$$

GRNN consists of four layers: input, hidden, summation, and output layers. The greatest advantage of GRNN is the training speed. Meanwhile, it is relatively insensitive to outliers (wild points). Training a GRNN actually consists mostly of copying training cases into the network, and so is as close to instantaneous as can be expected. The greatest disadvantage is network size: a GRNN network actually contains the entire set of training cases, and is therefore space-consuming, requiring more memory space to store the model. Relative literatures are [192,250-255].

Support vector machine (SVM)

SVM, developed by Vapnik [256,257] as a novel type of machine learning method, is gaining popularity due to its many attractive features and promising empirical performance. Originally, SVM was developed for pattern recognition problems. After that, SVM it was applied to regression by the introduction of an alternative loss function and results appear to be very encouraging [258]. As a developing method, new types of SVM are coming in on the stage of QSAR/QSPR, such as Least Square Support Vector Machine (LS-SVM), Grid Search Support Vector Machine (GS-SVM), Potential Support Vector Machine (P-SVM) and Genetic Algorithms Support Vector Machine (GA-SVM). Here, we only choose LS-SVM, the most commonly used one, as an example to provide a description.

Least square support vector machine (LS-SVM)

The LS-SVM, which was a modified SVM algorithm, was proposed by Suykens et al. [259] and used to build the nonlinear model. Here, we only briefly describe the main idea of LS-SVM for function estimation. In principle, LS-SVM always fits a linear relation (y = wx+b) between the regressors (x) and the dependent variable (y). The best relation can be obtained by minimizing the cost function (Q) containing a penalized regression error term:

$$QLS\text{-}\ SVM=12wTw+\gamma\sum k=1Nek2\text{-}(58)$$

subject to:0020

$$yk=wT\phi(xk)+b+ek,\ k=1,\ \ldots,\ N\text{-}(59)$$

where $\phi : Rn \rightarrow Rm$ is the feature map mapping the input space to a usually high-dimensional feature space, γ is the relative weight of the error term, and ek is error variables taking noisy data into accurate and avoiding poor generalization. LS-SVM considers this optimization problem to be a constrained optimization problem and uses a language function to solve it. By solving the Lagrange style of equation (58), the weight coefficients (w) can be written as:

$$w=\sum k=1N\alpha kxkTx+b \text{ with } \alpha k=2\gamma ek\text{-}(510)$$

By substituting it into the original regression line (y □ wx + b), the following result can be obtained:

$$y=\sum k=1N\alpha kxkTx+b\text{-}(511)$$

It can be seen that the Lagrange multipliers can be defined as:

$$\alpha k=(xkTx+(2\gamma)\text{-}1(yk\text{-}b))$$

Finding these Lagrange multipliers is very simple, as opposed to with the SVM approach, in which a more difficult relation has to be solved to obtain these values. In addition, it easily allows for a nonlinear regression as an extension of the linear approach by

introducing the kernel function. This leads to the following nonlinear regression function:

$$f(x)=\sum k=1N\alpha kK(x,xk)+b-(512)$$

where K(x, xk) is the kernel function. The value is equal to the inner product of two vectors x and xk in the feature space Φ(x) and Φ(xk); that is, K(x, xk)) = Φ(x)TΦ(xk). The choices of a kernel and its specific parameters together with γ have to be tuned by the user. The radial basis function (RBF) kernel K(x, xk) = exp(-||xk - x||2/σ2) is commonly used, and then leave-one-out (LOO) cross-validation is employed to tune the optimized values of the two parameters γ and σ. LS-SVM is a simplification of traditional support vector machine (SVM). It encompasses similar advantages with SVM and its own additional advantages. It only requires solving a set of linear equations (linear programming), which is much easier and computationally simpler than nonlinear equations (quadratic programming) employed by traditional SVM. Therefore, LS-SVM is faster than traditional SVM in treating the same work. The related literature is presented in [183,259-263].

Gene expression programming (GEP)

Gene expression programming was invented by Ferreira in 1999 and was developed from genetic algorithms and genetic programming (GP). GEP uses the same kind of diagram representation of GP, but the entities evolved by GEP (expression trees) are the expression of a genome. GEP is simpler than cellular gene progression. It mainly includes two sides: the chromosomes and the expression trees (ETs). The process of information of gene code and translation is very simple, such as a one-to-one relationship between the symbols of the chromosome and the functions or terminals they represent. The rules of GEP determine the spatial organization of the functions and terminals in the ETs and the type of interaction between sub-ETs. Therefore, the language of the genes and the ETs represents the language of GEP.

The GEP chromosomes, expression trees (ETs), and the mapping mechanism: Each chromosome in GEP is a character string of fixed length, which can be composed of gene from the function set or the terminal set. Using the elements {+, -,*, /, Q} as the function set and {a, b, c, d} as the terminal set, the following is an example of GEP chromosome of length eight:

$$\{01234567Q*+-abcd\}$$

where Q denotes the square-root function; and a, b, c, d are variable (or attribute) names. The above is referred to as Karva notation or K-expression. A K-expression can be mapped into an ET following a width-first procedure. A branch of the ET stops growing when the last node in this branch is a terminal. The conversion of an ET into a K-expression is also very straightforward and can be accomplished by recording the nodes from left to right in each layer of the ET in a top-down fashion to form the string:

$$y= (a-b)*(c+d)-(513)$$

Description of the GEP algorithm: The purpose of symbolic regression or function finding is to find an expression that can give a good explanation for the dependent variable. The first step is to choose the fitness function. Mathematically, the fitness fi of an individual program i is expressed by the equation:

$$fi=\sum j=1n(R-|P(ij)-TjTj.100|)-(514)$$

where R is the selection range, P(ij) is the value predicted by the individual program i for fitness case j (out of n fitness cases), and Tj is the target value for fitness case j. Note that the absolute value term corresponds to the relative error. This term is what is called the precision and if the error is smaller than or is equal to the precision then the error becomes zero. Thus, for a good match the absolute value term is zero and fi = fmax = nR. For some function finding problems it is important to evolve a model that performs well for all fitness cases within a certain relative error (the precision) of the correct value:

$$f(i,j)=\{1,E(i,j)\leq p0,-(515)$$

where p is the precision and E(ij) is the relative error of an individual program i for the fitting case j.

The E(ij) is given by:

$$E(ij)=|P(ij)-TjTj\bullet100|-(516)$$

The second step consists of choosing the set of terminals T and the set of functions F to create the chromosomes. In this problem, the terminal set consists obviously of the independent variable, i.e., T = {a}. The third step is to choose the chromosomal architecture, i.e., the length of the head and the number of genes. The fourth major step is to choose the linking function. The last major step is to choose the set of genetic operators that cause variation and their rates. These processes are repeated for a pre-specified number of generations until a solution is obtained. In the GEP, the individuals are often selected and copied into the next generation based on their fitness, as determined by roulette-wheel sampling with elitism, which guarantees the survival and cloning of the best individual to the next generation. The variation in the population is introduced by applying one or more genetic operators to select chromosomes, including crossover, mutation, and rotation. The process begins with the random generation of the chromosomes of the initial population. The chromosomes are expressed and the fitness of each individual is evaluated. The individuals are selected according to the fitness to reproduce with modification, leaving progeny with new traits. The individuals of this new generation are, in their turn, subjected to the same developmental process: expression of the genomes, confrontation of the selection environment, and reproduction with modification. To evaluate the ability of the GEP, the correlation coefficient (R) was introduced as:

$$Ci=Cov(T,P)\sigma t\sigma p\text{-}(517)$$

Where Cov(T, P) is the covariance of the target and model outputs. σt and σp are the corresponding standard deviations. GEP is the newest chemometrics method, and Si et al. [63,171,59-266] have applied this method to QSAR studies for the first time. The results from their studies are satisfactory and show a promising use in the nonlinear structure-activity/property relationship correlation area, but GEP is congenitally defective as far as reproducibility of the predicted results is concerned, and always deduces too complex equations. This means a higher requirement for a user who is involved with GEP. GEP is now commercialized in the software Automatic Problem Solver 3.0 or GeneXproTools 4.0 [267].

Project pursuit regression (PPR)

PPR, which was developed by Friedman and Stuetzle [115], is a powerful tool for seeking the interesting projections from high-dimensional data into lower dimensional space by means of linear projections. Therefore, it can overcome the curse of dimensionality because it relies on estimation in at most trivariate settings. Friedman and Stuetzle's concept of PPR avoided many difficulties experienced with other existing nonparametric regression procedures. It does not split the predictor space into two regions, thereby allowing, when necessary, more complex models. In addition, interactions of predictor variables are directly considered because linear combinations of the predictors are modeled with general smooth functions. Another significant property of PPR is that the results of each interaction can be depicted graphically. The graphical output can be used to modify the major parameters of the procedure: the average smoother bandwidth and the terminal threshold. A brief description about PPR is presented here. Given the (k × n) data matrix X, where k is the number of observed variables and n is the number of units, and an m-dimensional orthonormal matrix A (m × k), the (m × n) matrix Y = AX represents the coordinates of the projected data onto the m-dimensional (m < k) space spanned by the rows of A. Because such projections are infinite, it is important to have a technique to pursue a finite sequence of projections that can reveal the most interesting structures of the data. Projection pursuit (PP) is such a powerful tool that combines both ideas of projection and pursuit. In a typical regression problem, PPR aims to approximate the regression pursuit function f(x) by a finite sum of ridge functions with suitable choices of αi and gi:

$$g(p)(x)=\sum i=1pgi(\alpha iTx)\text{-}(518)$$

where αi values are m × n orthonormal matrices and p is the number of ridge functions. In recent QSAR/QSPR studies [321,337,424-429] PPR was employed as a regression method and always resulted in the best final models. This indicates that PPR is a promising regression method in QSAR/QSPR studies, especially when the correlation between descriptors and activities or properties is nonlinear.

Local lazy regression (LLR)

Most QSAR/QSPR models often capture the global structure-activity/property trends which are present in a whole dataset. In many cases, there may be groups of molecules which exhibit a specific set of features which relate to their activity or property. Such a major feature can be said to represent a local structure activity/property relationship. Traditional models may not recognize such local relationships. LLR is an excellent approach which extracts a prediction by locally interpolating the neighboring examples of the query which are considered relevant according to a distance measure, rather than considering the whole dataset. That will cause the basic core of this approach which is a simple assumption that similar compounds have similar activities or properties; that is, the activities or properties of molecules will change concurrently with the changes in the chemical structure. For one or more query points, "lazy" estimates the value of an unknown multivariate function on the basis of a set of possibly noisy samples of the function itself. Each sample is an input/output pair where the input is a vector and the output is a number. For each query point, the estimation of the input is obtained by combining different local models. Local models considered for combination by lazy are polynomials of zeroth, first, and second degree that fit a set of samples in the neighborhood of the query point. The neighbors are selected according to either the "Manhattan" or the "Euclidean" distance. It is possible to assign weights to the different directions of the input domain for modifying their importance in computation of the distance. The number of neighbors used for identifying local models is automatically adjusted on a query-by-query basis through a leave-one-out validation of models, each fitting a different number of neighbors. The local models are identified by using the recursive least-squares algorithm, and the leave-one-out cross-validation is obtained through the PRESS statistic. We assumed that there existed a small region around x which is linear between the dependent variable and the predictor variables. Then we determined the points around XNN(x) and build a regression model with only the points in NN(x) using the least-squares method and minimized the squared residuals for the region using this model. So we do not need to build multiple models for each point in the training set beforehand. In essence, when faced with a query point, the approach builds a representative predictive model. Hence, this approach is termed local LLR, which is described below:

$$\beta x=(XNN(X)TXNN(x))\text{-}1XNN(X)TYNN(x)\text{-}(519)$$

where XNN(x) was a matrix of independent variables, YNN(x) the column vectors representing the dependent variables, for the molecules in the neighborhood of the query point. βx was the column vectors of regression coefficients. One of the main advantages of LLR is the fact that no a priori model needs to be built. This makes it suitable for large data sets, where using all of the observations can normally be time-consuming and even lead to over-fitting. At the same time, because it builds a regression model for each query point, one cannot extract meaningful structure-activity trends for the data set as a whole. That is, the focus of LLR is on predictive ability, rather than interpretability. Like every method, the lazy approach has a number of shortcomings. First, as all of the computations are done at query time, the determination of the local neighborhood must be efficient. Second, uncorrelated features might result in errors

in the identification of near neighbors. Finally, it is nontrivial to integrate feature selection in this framework. LLR is generally used to develop linear models for data sets in which the global structure-activity/property relationship is nonlinear in nature. However, as a new arising method in the field of QSAR/QSPR, LLR is not used extensively, with only a few relevant studies shown in references [294,295-307]. It is expected that more application studies involving LLR will appear in future QSAR/QSPR analyses.

Theory, Implementation and Applications of Non-stationary Gabor Frames

Signal analysis with classical Gabor frames leads to a fixed time-frequency resolution over the whole time-frequency plane. To overcome the limitations imposed by this rigidity, we propose an extension of Gabor theory that leads to the construction of frames with time-frequency resolution changing over time or frequency. We describe the construction of the resulting nonstationary Gabor frames and give the explicit formula for the canonical dual frame for a particular case, the painless case. We show that wavelet transforms, constant-Q transforms and more general filter banks may be modeled in the framework of nonstationary Gabor frames. Further, we present the results in the finite-dimensional case, which provides a method for implementing the above-mentioned transforms with perfect reconstruction. Finally, we elaborate on two applications of nonstationary Gabor frames in audio signal processing, namely a method for automatic adaptation to transients and an algorithm for an invertible constant-Q transform. Redundant short-time Fourier methods, also known as Gabor analysis [216], are widely used in signal processing applications. The basic idea is the analysis of a signal f by consideration of the projections $\langle f,g_{\tau,\omega}\rangle$ of f onto time-frequency atoms $g_{\tau,\omega}$. The $g_{\tau,\omega}$ are obtained by translation of a unique prototype function over time and frequency: $g_{\tau,\omega}(t)=g(t-\tau)e^{2\pi it\omega}$. This classical construction leads to a signal decomposition with fixed time-frequency resolution over the whole time-frequency plane. The restriction to a fixed resolution is often undesirable in processing signals with variable time-frequency characteristics. Alternative decompositions have been introduced to overcome this deficit, e.g. the wavelet transform [217], the constant-Q transform (CQT) [218] or decompositions using filter banks [219], in particular based on perceptive frequency scales [220]. Adaptation over time is considered in approaches such as modulated lapped transforms [221], adapted local trigonometric transforms or (time-varying) wavelet packets. Most of the cited work achieves flexible tilings of the time-frequency plane, but efficient reconstruction from signal-adaptive, overcomplete time-frequency transforms is rarely addressed [222,223]. One exception is a recent approach in [224], which is in fact a special case of the more general model considered in the present paper. The wealth of existing approaches to fast adaptive transforms underlines the need for flexibility arising from many applications. On the other hand, the introduction of flexibility in a transform that is based on accurate mathematical modeling causes technical complications that are not always easy to overcome. We introduce an approach to fast adaptive time-frequency transforms that is based on a generalization of painless nonorthogonal expansions [225]. It allows for adaptivity of the analysis windows and the sampling points. Since the resulting frames locally resemble classical Gabor frames and share some of their structure, they are called nonstationary Gabor frames. The corresponding transform is likewise referred to as nonstationary Gabor transform (NSGT).

The central feature of painless expansions is the diagonality of the frame operator associated with the proposed analysis system. This idea is used here to yield painless nonstationary Gabor frames and will allow for both mathematical accuracy in the sense of perfect reconstruction (the frame operator is invertible) and numerical feasibility by means of an FFT-based implementation. The construction of painless nonstationary Gabor frames relies on three intuitively accessible properties of the windows and time-frequency shift parameters used.

1. The signal f of interest is localized at time- (or frequency-) positions n by means of multiplication with a compactly supported (or limited bandwidth, respectively) window function g_n.

2. The Fourier transform is applied on the localized pieces $f\cdot g_n$. The resulting spectra are sampled densely enough in order to perfectly reconstruct $f\cdot g_n$ from these samples.

3. Adjacent windows overlap to avoid loss of information. At the same time, unnecessary overlap is undesirable. In other words, we assume that $0<A\leq\sum_{n\in Z}|g_n(t)|^2\leq B<\infty$, a.e., for some positive A and B.

We will show that these requirements lead to invertibility of the frame operator and therefore to perfect reconstruction. Moreover, the frame operator is diagonal and its inversion is straight-forward. Further, the canonical dual frame has the same structure as the original one. Because of these pleasant consequences following from the three above-mentioned requirements, the frames satisfying all of them will be called painless nonstationary Gabor frames and we refer to this situation as the painless case. Under application of a Fourier transform to the signal of interest, our approach leads to adaptivity in either time or frequency. The concept of this paper relies on ideas introduced in [226], and presented in [227]. In the present paper all formal proofs are given, the link to frame theory is provided, the possibility to represent other analysis/synthesis systems with this approach is established, the numerical issues are investigated and several applications are presented.

The rest of the article is organized as follows. We fix notation and review preliminary results from Gabor and frame theory in Section 2. Section 3 introduces the construction of (painless) nonstationary Gabor frames in detail and provides a proof for the frame property under the given conditions. The calculation of the dual or tight frames is also explicitly given for systems adaptive in time or frequency, respectively. Section 4 then establishes the details of implementation in a discrete and real-life setting and provides examples together with a comparison of numerical efficiency with existing approaches. We conclude, in Section 5 with a summary and a brief outlook on future work.

In the sense of reproducible research, [228], we provide all algorithms and scripts to reproduce the results in this paper at the webpage http://univie.ac.at/nonstatgab/.

Please note that a nonstationary Gabor transform is also included in the Linear Time Frequency Analysis Toolbox (LTFAT) v.1.0 [285,286], a Matlab/Octave toolbox, which is freely available at http://ltfat.sourceforge.net/.

Preliminaries

For an integrable function f, i.e. f∈L1(R), we denote its Fourier transform Ff(ξ)=f^(ξ)=∫Rf(t)e-2πiξtdt, with the usual extension to L2(R), the space of square-integrable functions from R to C. The convolution of two functions f,g∈L1(R) is the function f∗g defined by (f∗g)(t)=∫Rf(x)g(t-x)dx, again with the usual extension to L2(R). It follows that F(f∗g)=f^·g^. We use the notation f(t)≃g(t) if there exist constants C1,C2>0, such that C1g(t)≤f(t)≤C2g(t) for all t.

Frame theory: We now give a short summary of frame theory on Hilbert spaces, first introduced in [231]. A thorough discussion can be found in [232,233]. A sequence (ψl)l∈I in the Hilbert space H is called a frame, if there exist positive constants A and B (called lower and upper frame bounds, respectively) such that

$$A||f||2\leq\sum l\in I|\langle f,\psi l\rangle|2\leq B||f||2\forall f\in H,\text{-}(61)$$

i.e. ∑l∈I|⟨f,ψl⟩|2≃||f||2. If A=B, then (ψl)l∈I is a tight frame. By C:H→ℓ2, we denote the analysis operator defined by (Cf)l=⟨f,ψl⟩. The adjoint of C∗ of C is the synthesis operator C∗(cl)=∑lclψl. The frame operator is Sf=C∗Cf=∑l⟨f,ψl⟩ψl, hence ⟨Sf,f⟩=||Cf||ℓ22.

The boundedness and invertibility of S is equivalent to the existence of frame bounds 0<A,B<∞ in the frame inequality (1), as well as to the existence of dual frames, which can be used for reconstruction. In particular, the canonical dual frame (ψ~l), is found by applying the inverse of S to the original frame elements, i.e. ψ~l=S-1ψl for all l. For all f∈H we then have the following reconstruction formulas:

$$f=\sum l\langle f,\psi l\rangle\psi\tilde{}l=\sum l\langle f,\psi\tilde{}l\rangle\psi l.$$

For tight frames, the frame operator reduces to S=AI, where I denotes the identity operator, and therefore S-1=1AI. The canonical tight frame (ψ°l) is obtained by applying S-12 to the frame elements, i.e. ψ°l=S-12ψl for all l.

Gabor theory: Recall that for any nonzero function g∈L2(R) (the window), the short-time Fourier transform (STFT) of a signal f∈L2(R) is defined as Vg(f)(τ,ω)=⟨f,MωTτg⟩, using the translation operator Tτf(t)=f(t-τ) and the modulation operator Mωf(t)=f(t) e2πiωt. In L2(R), we have

Vg(f)(τ,ω)=∫Rf(t)g(t-τ)¯e-2πiωtdt.

For a non-zero window function g and parameters a,b>0, the set of time-frequency shifts of g

G(g,a,b)={MbmTang:m,n∈Z}

is called a Gabor system [19]. Moreover, if G(g,a,b) is a frame, it is called a Gabor frame and the associated frame operator is denoted by Sg,a,b. In the succeeding sections, where the dependence of the frame operator on the window g and the parameters a,b is clear, we simply denote the frame operator by S. Note that the Gabor analysis coefficients are sampling points of the STFT of f with window g at the time-frequency points (an,bm), i.e. Vg(f)(an,bm)={⟨f,MbmTang⟩}m,n.

A central property of Gabor frames is the fact that the dual frame of a Gabor frame is again a Gabor frame, generated by the dual window g~=S-1g and the same lattice, i.e. the set of time-frequency points {(an,bm)|m,n∈Z}. Note that the property that the dual system is again a system with the same structure is a particular property of Gabor frames, shared by nonstationary Gabor frames in the painless setting considered in the present paper. For a more detailed introduction to Gabor analysis, see [216,234]. In the finite discrete case, we take the Hilbert space H to be CL. For a good introduction to Gabor analysis in this setting, see [235]. We shall restrict the lattice parameters a&b to factors of L such that the numbers N=La and M=Lb are integers. We regard all vectors as periodic, so discrete translation is a cyclic operator. Therefore, the discretization

Tnx=(xL-n,xL-n+1,…,x0,x1,…,xL-n-1)

and

Mmx=(x0·WL0,x1·WL1·m,…,xL-1·WL(L-1)m)

with WL=e2πiL, respectively. We will consider the Gabor system

G(g,a,b)={MbmTang:n=0,…,N-1;m=0,…,M-1},

which is a collection of M·N vectors in CL. Obviously, to fulfill the frame conditions (1), we need at least M·N≥L.

Wavelet theory: As we will see below, nonstationary Gabor frames may be used to construct wavelet frames. We briefly sketch the continuous wavelet transform. Let ψ∈L2(R) and (α,β)∈R+∗×R. We define the wavelet system by

$$\psi\alpha,\beta(t)=1\alpha\psi(t-\beta\alpha)=T\beta D\alpha\psi,\text{-}(62)$$

where Dα denotes the dilation operator given by Dαf(t)=1αf(tα).

The wavelet transform is then defined as

$$W\psi f(\alpha,\beta)=\langle f,T\beta D\alpha\psi\rangle=(f*D\alpha I\psi^{-})(\beta),-(63)$$

where I denotes the involution Ig(t)=g(-t).

If ψ is localized around τ0, then ψα,β(t) is centered at α·τ0+β. The frequency center is at η/α, where η is the center of ψ^.

Construction of non-stationary Gabor frame

Resolution changing over time: As opposed to standard Gabor analysis, where time translation is used to generate atoms, the setting of nonstationary Gabor frames allows for changing, hence adaptive, windows in different time positions. Then, for each time position, we build atoms by regular frequency modulation. Using a set of functions {gn}n∈Z in L2(R) and frequency sampling step bn, for m∈Z and n∈Z, we define atoms of the form:

gm,n(t)=gn(t)e2πimbnt=Mmbngn(t), implicitly assuming that the functions gn are well-localized and centered around time-points an. This is similar to the standard Gabor scheme, however, with the possibility to vary the window gn for each position an. Thus, sampling of the time-frequency plane is done on a grid which is irregular over time, but regular over frequency at each temporal position. Note that some results exist in Gabor theory for semi-regular sampling grids, as for example in [294]. Our study uses a more general setting, as the sampling grid is in general not separable and, more importantly, the window can evolve over time. To get a first idea of the effect of nonstationary Gabor frames, regular Gabor transforms and a nonstationary Gabor transform of the same signal. Note that the NSGT was adapted to transients and the components are well-resolved.

Example of a sampling grid of the time-frequency plane when building decomposition with time-frequency resolution evolving over time

Glockenspiel (Example 1) Gabor representations with short window (11.6ms), resp. long window (185.8ms)

Glockenspiel (Example 1) Regular Gabor representation with a Hann window of 58 ms length and a nonstationary Gabor representation using Hann windows of varying length.

In the current situation, the analysis coefficients may be written as

cm,n=⟨f,Mmbngn⟩=(f·gn¯)^(mbn),m,n∈Z.

Remark 1

If we set gn(t)=g(t-na) for a fixed time-constant a and bn=b for all n, we obtain the case of classical painless non-orthogonal expansions for regular Gabor systems introduced in [225].

Resolution changing over frequency: An analog construction in the frequency domain leads to irregular sampling over frequency, together with windows featuring adaptive bandwidth. Then, sampling is regular over time. An example of the sampling grid in such a case is given in Equation 4.

Example of a sampling grid of the time-frequency plane when building decomposition with time-frequency resolution changing over frequency.

In this case, we introduce a family of functions {hm}m∈Z of L2(R), and for m∈Z and n∈Z, we define atoms of the form:

$$hm,n(t)=hm(t-nam).-(64)$$

Therefore hm,n^(ν)=hm^(ν)·e-2πinamν and the analysis coefficients may be written as

cm,n=⟨f,hm,n⟩=⟨f^,F(Tnamhm)⟩=F-1(f^·hm^¯)(nam).

Hence, the situation is completely analog to the one described in the previous section, up to a Fourier transform. In practice we will choose each function hm as a well localized band-pass function with center frequency bn.

Link between nonstationary Gabor frames, wavelet frames and filter banks: To obtain wavelet frames, the wavelet transform in (217) is sampled at sampling points (βn,αm). A typical discretization scheme [227] is (nβ0,α0m). Then, the frame elements are ψm,n(t)=Tnβ0Dα0mψ(t). Comparing this expression to (4) and setting hm=Dα0mψ and am=β0, we see that a wavelet frame with this discretization scheme corresponds to a nonstationary Gabor transform.

Another possibility for sampling the continuous wavelet transform [217] uses α=α0m and β=nβ0α0m. Again, we obtain a correspondence to nonstationary Gabor frames by setting hm=Dα0mψ and am=β0·α0m.

Beyond the setting of wavelets, any filter bank [227], even with non-constant down-sampling factors Dm, can be written as a nonstationary Gabor frame. A filter bank is a set of time-invariant, linear filters hm, i.e. Fourier multipliers. The response of a filter bank for the signal f and sampling period T0 is given (in the continuous case) by

cm,n=(f∗hm)(nDmT0)=∫Rf(t)hm(nDmT0-t)dt=⟨f,hm,n⟩,

where hm,n(t)=h(nDmT0-t)¯. Setting hm=Ihm¯ and choosing am=DmT0 this construction is realized with nonstationary Gabor frames using [219]. If the filters are band-limited and the down-sampling factors are small enough, then the conditions for the painless case are met and the corresponding reconstruction procedure can be applied.

Invertibility of the frame operator and reconstruction: In this central section, we give the precise conditions under which painless nonstationary Gabor frames are constructed. The first two basic conditions, namely compactly supported windows and sufficiently dense frequency sampling points, lead to diagonality of the associated frame operator S, as defined in Section 2.1. The third condition, the controlled overlap of adjacent windows, then leads to boundedness and invertibility of S. The following theorem generalizes the results given for the classical case of painless non-orthogonal expansions [225,234].

Theorem 1

For every n∈Z, let the function gn∈L2(R) be compactly supported with supp (gn) ⊆[cn,dn] and let bn be chosen such that dn-cn≤1bn . Then the frame operator

S:f↦∑m,n⟨f,gm,n⟩gm,n of the system

gm,n(t)=gn(t)e2πimbnt,m∈Zandn∈Z,

is given by a multiplication operator of the form

Sf(t)=(∑n1bn|gn(t)|2)f(t).

Proof

Note that,

⟨Sf,f⟩=∑n∑m|∫Rf(t)gn(t)¯e-2πimbntdt|2=∑n∑m|∫cndnf(t)gn(t)¯e-2πimbntdt|2,

due to the compact support property of the windows gn. Let In=[cn,cn+bn-1] for all n and χI denote the characteristic function of the interval I. Taking into account the compact support of gn again, it is obvious that

fgn¯=χIn∑lTlbn-1(fgn¯),

with the bn-1-periodic function ∑lTlbn-1(fgn¯). Hence, with Wm,n(t)=e-2πimbnt,

|∫cndnf(t)gn(t)¯Wm,n(t)dt|2=|∫Inf(t)gn(t)¯Wm,n(t)dt|2,=|⟨fgn¯,Wm,n⟩L2(In)|2

and applying Parseval's identity to the sum over m yields

⟨Sf,f⟩=∑n∑m|⟨fgn¯,Wm,n⟩L2(In)|2=∑n1bn||fgn¯||2=⟨∑n1bn|gn|2f,f⟩.

While in general, the inversion of S can be numerically unfeasible, in the special case described in Theorem 1, the invertibility of the frame operator is easy to check and inversion is a simple multiplication.

Corollary 1

Under the conditions given in Theorem 1, the system of functions gm,n forms a frame for L2(R) if and only if ∑n1bn|gn(t)|2≃1 . In this case, the canonical dual frame elements are given by

g˜m,n(t)=gn(t)∑l1bl|gl(t)|2e2πimbnt,-(65)

and the associated canonical tight frame elements can be calculated as:

g°m,n(t)=gn(t)∑l1bl|gl(t)|2e2πimbnt.

Remark 2

The optimal lower and upper frame bounds are explicitly given by Aopt=essinf∑n1bn|gn(t)|2 and Bopt=esssup∑n1bn|gn(t)|2.

We next state the results of Theorem 1 and Corollary 1 in the Fourier domain. This is the basis for adaptation over frequency.

Corollary 2

For every m∈Z, let the function hm be band-limited to supp (hm^)=[cm,dm] and let am be chosen such that dn-cn≤1am . Then the frame operator of the system

hm,n(t)=hm(t-nam),m∈Z,n∈Z

is given by a convolution operator of the form

⟨Sf,f⟩=⟨F-1(∑m1am|hm^|2)∗f,f⟩-(66)

for f∈L2(R) . Hence, the system of functions hm, n forms a frame of L2(R) if and only if ∀ν∈R, ∑m1am|hm^(ν)|2≃1 . The elements of the canonical dual frame are given by

h˜m,n(t)=TnamF-1(hm^∑l1al|hl^|2)(t)-(67)

and the canonical tight frame is given by

h°m,n(t)=TnamF-1(hm^∑l1al|hl^|2)(t).-(8)

Proof

We deduce the form of the frame operator in the current setting from the proof of Theorem 1 by setting

⟨Sf,f⟩=⟨Sf^,f^⟩=∑m,n|⟨f^,hm,n^⟩|2 and the rest of the corollary is equivalent to Corollary 1.

Remark 3

Classical Gabor frames are intimately related to modulation spaces, see [237] for an extensive discussion and relevant references. The characterization of modulation spaces depends on the joint time-frequency localization of the analysis window. Painless nonstationary Gabor frames characterize modulation spaces, if, in addition to compactness in one domain (time or frequency), the windows gk exhibit a uniform decay in the sense of time-frequency molecules, see [235], Theorem [22], i.e., letting ξ=(ak,l/bk),k,l∈Z, we require |Vφgk(z)|≤C(1+|z-ξ|)-r for some r>2. Then, the corresponding frame operator is invertible on all modulation spaces Mp,1≤p≤∞, and the ℓp-norm of the corresponding coefficient sequence is equivalent to the modulation space norm.

Remark 4

As mentioned in Section 3.2.1 the NSGT is linked to wavelet frames. In the painless case it is possible to construct a dual sequence which has the same structure. For wavelets this is also possible, see e.g. [239,240], where non-canonical duals are constructed. In a similar way as modulation spaces are linked to the Gabor transform, Besov spaces are related to wavelet systems; see e.g. [241]. Also, Sobolev spaces can be linked to the wavelet transform. Nonstationary Gabor frames can also be used to characterize Besov and Sobolev spaces, with some additional assumptions. Details will be reported elsewhere.

Discrete finite nonstationary Gabor frames

Discrete, time-adaptive Gabor transform: For the practical implementation, the equivalent theory may be developed in a finite discrete setting using the Hilbert space CL. Since this is largely straight-forward from simple matrix multiplication, we only state the main result. Given a set of functions {gn}n∈{0,...,N-1}, a set of integers (number of frequency samples for each time position) {Mn} n∈{0,...,N-1} associated with the set of real values {bn=LMn}n∈{0,...,N-1}, the discrete, nonstationary Gabor system is given by

gm,n[k]=gn[k]·e2πimbnkL=gn[k]·WLmbnk

for n=0,...,N-1,m=0,...,Mn-1 and all k=0,...,L-1. Note that in practice, gm,n[k] will have zero-values for most k, allowing for efficient FFT-implementation: since Mn=Lbn, we have gm,n[k]=gn[k]·e2πimkMn and the nonstationary Gabor coefficients are given by an FFT of length Mn for each gn.

The number of elements of {gm,n} is P=∑n=0N-1Mn. Let G be the L×P matrix such that its p-th column is gm,n, for p=m+∑k=0n-1Mk.

Corollary 3

The frame operator S=G·G∗ is an L×L matrix with entries:

Sk, j=∑n∈N(k-j)Mngn[k]gn[j]¯

where Np={n∈[0,N-1]|p=0modMn} for p∈[-L,L] . Therefore, if appropriate support conditions are met, S is a diagonal matrix.

Numerical complexity: Assuming that the windows gn have support of length Ln, let M=maxn {Mn} be the maximum FFT-length. We consider the painless case where Ln≤Mn≤M. The number of operations is

1. Windowing: Ln operations for the n-th window.

2. FFT: O (Mn·log (Mn)) for the n-th window.

Then the number of operations for the discrete NSGT is

O(∑n=0N-1Mn·log(Mn)+Ln)=O(N·(Mlog(M)+M))=O(N·(Mlog(M))).

Similar to the regular Gabor case, the number of windows N will usually depend linearly on the signal length L while the maximum FFT-length M is assumed to be independent of L. In that case, the discrete NSGT is a linear cost algorithm.

For the construction of the dual windows in the painless case, the computation involves multiplication of the window functions by the inverse of the diagonal matrix S and results in O(2∑n=0N-1Ln)=O(N·M) operations. Finally, the inverse NSGT has numerical complexity O (N·(Mlog(M))), as in the NSGT, since it entails computing the IFFT of each coefficient vector, multiplying with the corresponding dual windows and evaluating the sum. Technical framework all subsequently presented simulations were done in MATLAB R2009b on a 2 Gigahertz Intel Core 2 Duo machine with 2 Gigabytes of RAM running Kubuntu 9.04. The CQTs were computed using the code published

with [30], available for free download at http://www.elec.qmul.ac.uk/people/anssik/cqt/.

The constant-Q nonstationary Gabor transform (CQ-NSGT) algorithms are available at http://univie.ac.at/nonstatgab/.

Application: automatic adaptation to transients: In real-life applications, NSGT has the potential to represent local signal characteristics, e.g. transient sound events, in a more appropriate way than pre-determined, regular transform schemes. Since the appropriateness of a representation depends on the specific application, any adaptation procedure must be designed specifically. For the implementation itself, however, two observations generally remain true: first, the general nonstationary framework needs to be restricted to a well defined set of choices. Second, some measure is needed to determine the most suitable of the possible choices [308]. For example, in the case of a sparsity measure, the most sparse representation will be chosen. To show that good results are achieved even when using quite simple adaptation methods, we describe a procedure suitable for signals consisting mainly of transient and sinusoidal components. The adaptation measure proposed is based on onset detection, i.e. estimating where transients occur in the signal. The transform setting is what we call scale frames: the analysis procedure uses a single window prototype and a countable set of dilations there of.

For evaluation, the representation quality is measured by comparison of the number of representation coefficients leading to certain root mean square (RMS) reconstruction errors, for both NSGT and regular Gabor transforms. The results are especially convincing for sparse music signals with high energy transient components. Other possible adaptation methods might be based on time-frequency concentration, sparsity or entropy measures [15,37,38].

Scale frames. In the following paragraphs, we propose a family of nonstationary Gabor frames that allows for exponential changes in time-frequency resolution along time positions. To avoid heavy notation and since the formalism necessary for the discrete, finite case could obscure the principal idea, we describe the continuous case construction. Suitable standard sampling then yields discrete, finite frames with equivalent characteristics. The basic idea is to build a sequence of windows gn from a single, continuous window prototype g with support on an interval of length 1 in such a way that the resulting gn satisfy Corollary 1. The window sequence will be unambiguously determined by a sequence of scales. Once this scale sequence is known, it is a simple task to choose modulation parameters bn satisfying the necessary conditions.

As a scale sequence, we allow any integer-valued sequence {sn}n∈Z such that |sn-sn-1|∈{0,1}, where the latter restriction is set in order to avoid sudden changes of window length. Then, gn is, up to translation, given by a dilation of the prototype g:

D2sn(g)(t)=2-sng(2-snt).

This implies that a change of scale from one time step to the next corresponds to the use of a window either half or twice as long. More precisely, for every time step n, set s=min{sn-1,sn} and fix an overlap of 2/3·2s, if sn≠sn-1 and 1/3·2s, if sn=sn-1. Explicitly,

gn=TnD2sn(g), with recursively defined time shift operators Tn given by

T0=T0,Tn={T2s5/6Tn-1,if sn≠sn-1T2s+1/3Tn-1,else.

Defining the time shifts in this manner, we achieve exactly the desired overlap as illustrated in equation 5.

Illustration of scale frame overlaps and time shifts: By construction, each gn has non-zero overlap with its neighbors gn-1 and gn+1 and at any point on the real line, at most two windows are non-zero. After performing a preliminary transient detection step, as explained before, the construction of the adapted frame reduces to the determination of a scale sequence.

In the subsequent Figures and experiments, we used the Hann window as the prototype, but other window choices are possible. The described concept can easily be generalized by admitting other overlap factors and scaling ratio than the ones specified above [309]. The parameters have to be chosen with some care, though. Otherwise, the resulting frames might be badly conditioned, with a big or even infinite condition number BA, caused by accumulation points for the time shifts or gaps between windows. A more detailed description of general and discrete scale frames is beyond the scope of this article and will be part of a future contribution.

Frame construction from a sequence of onsets. In this paragraph, we assume that the signals of interest are mainly comprised of transient and sinusoidal components, an assumption met, e.g. by piano music. The instant a piano key is hit corresponds to a percussive, transient sound event, directly followed by harmonic components, concentrated in frequency. An intuitive adaptation to signals of this type would use high time resolution at the positions of transients. This corresponds to applying minimal scale at the transients and steadily increasing the scale with the distance from the closest transient. The transients' positions can be determined, e.g. by the so-called onset detection procedures [39] which, if used carefully, work to a high degree of accuracy. Once the transient positions are known, the construction of a corresponding scale frame yields good nonstationary representations for sufficiently sparse signals.

Application of onset-based scale frames. We applied the procedure proposed above to various signals, mainly piano music [310]. For this presentation, we selected three examples, all of them sampled at 44.1 kHz and consisting of a single channel. Some more examples and corresponding results as well as the source sound files can be found on the associated web-page http://univie.ac.at/nonstatgab/.

- Example 1: the widely used Glockenspiel signal.
- Example 2: an excerpt from a solo jazz piano piece performed by Herbie Hancock, characterized by its calmness and varied

rhythmical pattern, resulting in irregularly spaced low-energy transients.

Hancock (Example 2). Regular and nonstationary Gabor representations

• Example 3: a short excerpt of György Ligeti's piano concert. With highly percussive onsets in the piano and Glockenspiel voices and some orchestral background, this is the most polyphonic of our examples.

Ligeti (Example 3). Regular and nonstationary Gabor representations

For comparison, the plots in Figures also show standard Gabor coefficients with comparable (average) window overlap. A Hann window of 2560 samples length was chosen for the computation of regular Gabor transforms. The comparison shows that for the three signals, the NSGT features a better concentration of transient energy than a regular Gabor transform, while keeping, or even improving, frequency resolution.

Efficiency in sparse reconstruction: The onset detection procedure and a subsequent scale frame analysis were applied, along with a regular Gabor decomposition, to the Glockenspiel and Ligeti signals. As a test of the representations' sparsity, the signals were synthesized from their corresponding coefficients, modified by hard thresholding followed by reconstruction using the canonical dual frame. Then the numbers of largest magnitude coefficients needed for a certain relative root mean square (RMS) reconstruction error for each representation were compared. The RMS error of a vector f and its reconstruction frec is given by

RMS (f, frec)=∑k=0L-1|f[k]-frec[k]|2∑k=0L-1|f[k]|2.

All transforms are of redundancy about 53. The results for NSGT and different regular Gabor transform schemes are listed in Figure 1. On the Glockenspiel signal the NSGT method performs vastly better than the ordinary Gabor transform. For Ligeti, the differences are not as significant, but still the NSGT-based procedure shows better overall results (Figure 1).

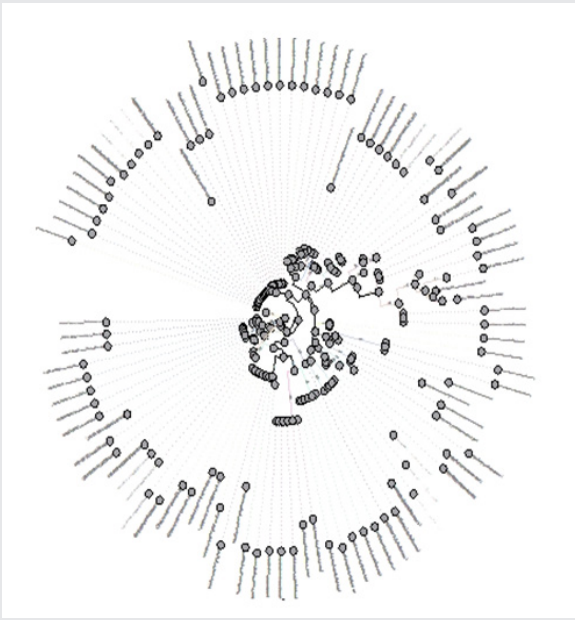


Figure 1

RMS error in sparse representations of Example 1,3. Parameters (in parentheses) are hop size and window length in the regular case (GT) or shortest window length and number of scales for the nonstationary case (NSGT). The values are estimated. Further experiments and a more exhaustive discussion of the parameters used in the experiments, can be found on the web-page http://univie.ac.at/nonstatgab/. Along with them, examples of regular and nonstationary reconstructions from a specified amount of coefficients can be found, so the reader might get a subjective impression of perceptive reconstruction quality. In conclusion, the experiments show that for real music signals, NSGT can provide a sparser representation than regular Gabor transforms, admitting reasonable reconstruction error.

Implementation of a discrete, frequency-adaptive Gabor transform: Since our construction of Gabor frames with adaptivity in the frequency domain relies on the fact that analysis windows hm possess compact band width; an FFT-based implementation is highly efficient. We take the input signal's Fourier transform and treat the procedure in complete analogy to the situation developed for time-adaptive transforms, i.e. hm,n[k]=Tnamhm[k] and hm,n^[j]=M-namhm^[j].

As observed in Section 3.2.1, we are able to obtain wavelet frames using Gabor frames that exhibit nonstationarity in the frequency domain. Moreover, we may design general transforms with flexible frequency resolution, such as a constant-Q transform. While various other adjustments (e.g. Mel- or Bark-scaled transforms) are feasible, we will focus our discussion on the constant-Q case. To the best knowledge of the authors, the approach to implement the constant-Q transform directly in the frequency domain by means of FFT is

new in audio processing.

Remark 5

Note that for real-valued signals the symmetry of their FFT can be exploited to further reduce the computational effort. We particularly refer to the LTFAT routines filterbankrealdual.m and filterbankrealtight.m.

Application: an invertible constant-Q transform: The constant-Q transform (CQT), introduced in [40], and transforms a time signal into the time-frequency domain, where the center frequencies of the frequency bins are geometrically spaced. Since the Q-factor (the ratio of the center frequencies to the window's bandwidth) is constant, the representation allows for a better frequency resolution at lower frequencies and a better time resolution at the higher frequencies [311]. This is sometimes preferable to the fixed resolution of the standard Gabor transform, for which the frequency bins are linearly spaced. In particular, this kind of resolution is often desired in the analysis of musical signals, since the transform can be set to co-incide the temperament, e.g. semitone or quarter tone, used in Western music.

The originally introduced constant-Q transform, however, is not invertible and is computationally more intensive than the DFT. A computationally more efficient approach was presented in the sequel [9]: for the nth time slice of the signal f, the coefficient vector cm,n, equal to inner product of the signal f with the time-limited window hm,n is computed in the Fourier side via ⟨f^,hm,n^⟩. This approximate computation takes advantage of the sparsity of the frequency domain kernel or spectral kernel. In contrast, we compute the coefficient vector for each frequency bin, making use of band-limited window functions.

Perfect reconstruction wavelet transforms with rational dilation factors were proposed in [41]. Since they are based on iterated filter banks, these methods are computationally too expensive for long, real-life signals, when high Q-factors, such as 12-96 bins per octave, are required.

In [36], Klapuri and Schörkhuber presented a computation of the CQT that shows improved efficiency and flexibility compared to the method proposed in [9], among others. However, the approximate inversion introduced in [36] still gives an RMS error of around 10-3. The lack of perfect invertibility prevents the convenient modification of CQT-coefficients with subsequent resynthesis required in complex music processing tasks such as masking or transposition. By allowing adaptive resolution in frequency, we can construct an invertible nonstationary Gabor transform with a constant Q-factor on the relevant frequency bins.Setting for the frame elements in the transform, we consider functions hm∈CL,m=1,...,M having center frequencies (in Hz) at ξm=ξmin2m-1B, as in the CQT. Here, B is the number of frequency bins per octave, and ξmin and ξmax are the desired minimum and maximum frequencies, respectively. In the experiments, we restrict ξmax to be less than the Nyquist frequency and there should exist an M∈N satisfying

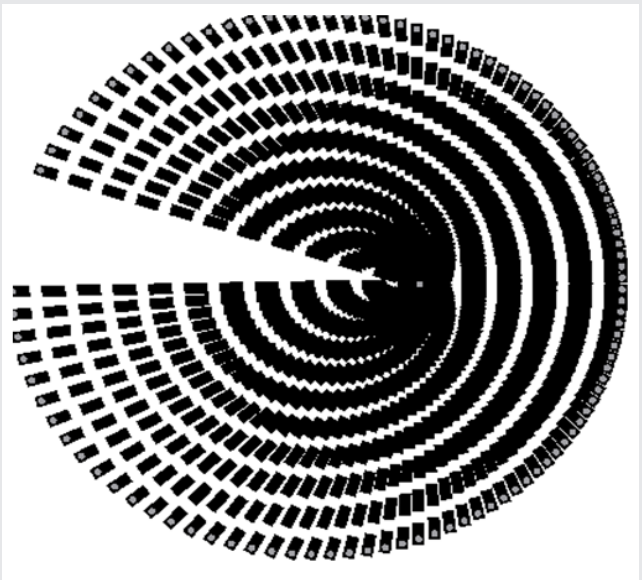


Figure 2: Exemplary sampling grid of the time–frequency plane for a constant-Q non-stationary Gabor system.

ξmax≤ξmin2M-1B<ξs/2,

where ξs denotes the sampling frequency. In this case, we take M=[Blog2(ξmax/ξmin)+1], where [z] is the smallest integer greater than or equal to z. While in the CQT no 0-frequency is present, the NSGT provides all necessary freedom to use additional center frequencies. Since the signals of interest are real-valued, we put filters at center frequencies beyond the Nyquist frequency in a symmetric manner. This results in the following values for the center frequencies:

ξm={0,m=0ξmin2m-1B,m=1,...,Mξs/2,m=M+1ξs-ξ2M+2-m,m=M+2,...,2M+1

For the corresponding bandwidth Ωm of hm, we set Ωm=ξm+1-ξm-1, for m=1,...,M, and Ω0=2ξ1=2ξmin. By construction, these result in a constant Q-factor Q=(21B-2-1B)-1 for m=2,...,M-1. And we can write each Ωm as follows:

Ωm={2ξmin,m=0ξ2,m=1,2M+1ξm/Q,m=2,...,M-1(ξs-2ξM-1)/2,m=M,M+2ξs-2ξM,m=M+1ξ2M+2-m/Q,m=M+3,...,2M.

If we use a Hann window hˆ, supported on [-1/2,1/2], then we can obtain each hm via hm^[j]=hˆ((jξsL-ξm)/Ωm), where j=0,...,L-1. Letting am≤ξsΩm, we define hm,n by their Fourier transform hm,n^=M-namhm^,n=0,...,⍰Lam⍰-1. Figure 2 illustrates the time-frequency sampling grid of the set-up, where the center frequencies are geometrically spaced and sampling points regularly spaced (Figure 2).

The support conditions on hm^ imply that the sum σ=∑m=02M+1Lam|hm^|2 is finite and bounded away from 0. From Section 3.3, the frame operator is therefore invertible and we can apply Corollary 2. Note that we consider the bandwidth to be the support of the window in frequency. This makes sense in the considered painless case. Very often, see e.g. (36), the bandwidth is taken as the width between the points, where the filter response drops to half of the maximum, i.e. the -3 dB-bandwidth. This definition would also make sense in a non-compactly supported case. For the chosen filters, Hann windows, the Q-factor considering the -3 dB-bandwidth is just double of the one considered above. We see in Figure 2 the standard Gabor transform spectrogram and the constant-Q NSGT spectrogram of the Glockenspiel signal, the latter being very similar to the CQT spectrogram obtained from the original algorithm [40] but with the additional property that the signal can be perfectly reconstructed from the coefficients. Figures 2&3 compare the standard Gabor transform spectrogram and the constant-Q NSGT spectrogram of two additional test signals, both sampled at 44.1 kHz.

Efficiency: The computation time of the nonstationary Gabor transform was found to be better than a recent fast CQT implementation (36), as seen in Figure 3. The two plots show mean values for computation time in seconds and the corresponding variance over 50 iterations, with varying window lengths and number of frequency bins, respectively. The outlier, drawn in gray, in Figure 3 (left) at the prime number 600 569 illustrates dependence of the current CQ-NSGT implementation on the signal length's prime factor structure, analogous to FFT.

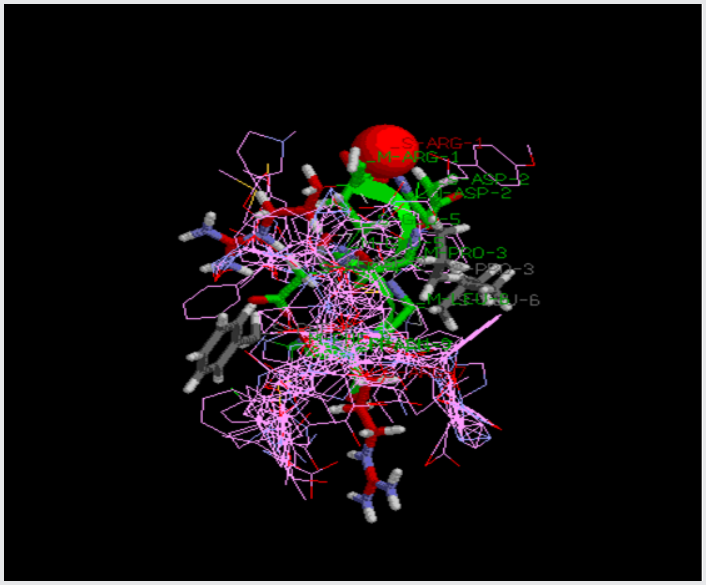


Figure 3: The PLS statistical results of CoMFA and CoMSIA model.

It is again reasonable to assume that the number of filters is bounded, independently of L, while the number of temporal points depends on L. As the role of M and N is switched in the assumption in Section 4.1.1 for the complexity, we arrive at a complexity of O (LlogL). This is also the complexity of the FFT of the whole signal [312]. So the overall complexity of the frequency-dependent nonstationary Gabor transform is O (LlogL). The advantage of the method in terms of computational efficiency thus decreases as longer signals are considered.

We note that at this point, since the windows used are band-limited, the current procedure is not suitable for real-time processing, despite its efficiency. The next step would be to process the incoming samples in a piecewise manner, using only a single family of frame elements for signals of arbitrary length. This entails working on finite, discrete parts of the given signal, thus considering the Fourier-transformed versions of vectors f·h∈CL, where h denotes some function of length L≪L. This window, together with the frame elements, will be designed to minimize undesired effects that stem from the cutting of the signal. Details of this piecewise processing, as well as a proposed variable-Q transform, will be further discussed in a future contribution.

xk+1 = xk - [∇e-diag(β1e1,β2e2,...,βnen)]-1(xk)e(xk),

where βi ∈ (0, 1) for i = 1, 2, ..., n. In their paper, only low dimension problems (two variables) are tested. In this experiment, we also give the numerical results of this method for large-scale nonlinear equations to compare with our proposed algorithm.

The parameters were chosen as r = 0.1, σ = 0.9, M = 12, ϵ = 10-4, and ϵmain = 10-5. In order to ensure the positive definiteness of Bk, in Step 4 of the main algorithm: if ykTsk>0, update Bk by (1.11), otherwise let Bk+1 = Bk. This program will also be stopped if the iteration number of main algorithm is larger than 200. Since the line search cannot always ensure these descent conditions dkTek<0 and dkT∇e(xk)ek<0, an uphill search direction may occur in numerical experiments. In this case, the line search rule maybe fails. In order to avoid this case, the stepsize αk will be accepted if the searching time is larger than six in the inner circle for the test problems.

In the sub-algorithm, the CG formula is used by the following Polak-Ribière-Polyak (PRP) method [61, 62]

dk={-ek+ekT(ek-ek-1)‖ek-1‖2dk-1if k≥1,-ekif k=0.-4.1

For the line search technique, (1.3) is used and the largest search number of times is ten,

where δ1 = δ2 = 10-7, and ϵk=1NI2

(NI is the iteration number). The sub-algorithm will also stopped if the iteration number is larger than 150. The iteration number, the function evaluations, and the CPU time of the sub-algorithm are added to the main algorithm for new method with CG. The meaning of the items of the columns of Table 1 is:

Numerical results

Dim: the dimension.

NI: the number of iterations.

NG: the number of function evaluations.

cpu time: the cpu time in seconds.

GF: the final norm function evaluations p(x) when the program is stopped.

GD: the final norm evaluations of search direction dk.

fails: fails to find the final values of p(x) when the program is stopped.

Table 1:

VYKSPNAYTLFS-ZINC14982928-1.pdb	-61.7401	-53.0282	-8.71191	0	19.3333
VYKSPNAYTLFS-ZINC14982847-1.pdb	-57.9674	-47.4746	-10.4927	0	18.5556
VYKSPNAYTLFS-ZINC14262823-0.pdb	-43.7743	-33.4255	-10.3489	0	18.4706
VYKSPNAYTLFS-ZINC14262821-0.pdb	-49.944	-36.3981	-13.5459	0	18.4706
VYKSPNAYTLFS-ZINC14134853-0.pdb	-62.9039	-48.9675	-13.9364	0	20.8
VYKSPNAYTLFS-ZINC13973961-0.pdb	-56.2728	-42.2728	-14	0	18.4706
VYKSPNAYTLFS-ZINC13906923-1.pdb	-57.7378	-48.8648	-8.87305	0	20
VYKSPNAYTLFS-ZINC13764320-0.pdb	-54.7926	-45.1902	-9.60237	0	22.4167
VYKSPNAYTLFS-ZINC13764156-0.pdb	-54.9457	-35.9318	-19.0139	0	21.6364
VYKSPNAYTLFS-ZINC13761951-1.pdb	-52.6511	-37.69	-14.961	0	19.6667
VYKSPNAYTLFS-ZINC13761947-1.pdb	-60.8774	-30.9049	-29.9725	0	18.8
VYKSPNAYTLFS-ZINC13600422-0.pdb	-59.0807	-40.673	-18.4077	0	20.8667
VYKSPNAYTLFS-ZINC13559074-1.pdb	-60.559	-46.559	-14	0	18.625
VYKSPNAYTLFS-ZINC13558381-1.pdb	-50.7924	-47.9386	-2.85376	0	18.4444
VYKSPNAYTLFS-ZINC13544268-0.pdb	-58.5835	-45.7292	-12.8543	0	20.7059
VYKSPNAYTLFS-ZINC13533175-1.pdb	-61.8268	-49.5924	-12.2343	0	18.5294
VYKSPNAYTLFS-ZINC13532312-1.pdb	-47.9765	-27.9069	-20.0697	0	25.375
VYKSPNAYTLFS-ZINC13529467-0.pdb	-62.4088	-43.8662	-18.5426	0	19.5294
VYKSPNAYTLFS-ZINC13529464-1.pdb	-66.7377	-47.3238	-19.4139	0	20.4706
VYKSPNAYTLFS-ZINC13529461-1.pdb	-54.1941	-47.1941	-7	0	20.125
VYKSPNAYTLFS-ZINC13529458-0.pdb	-60.3316	-49.8316	-10.5	0	19.5
VYKSPNAYTLFS-ZINC13526737-1.pdb	-52.1506	-35.2471	-16.9035	0	19.5714

VYKSPNAYTLFS-ZINC13526734-1.pdb	-59.1516	-36.3732	-22.7784	0	20.6429
VYKSPNAYTLFS-ZINC13526732-1.pdb	-56.5919	-35.1981	-21.3938	0	20.7857
VYKSPNAYTLFS-ZINC13526729-0.pdb	-61.4171	-33.948	-27.4691	0	21.4545
VYKSPNAYTLFS-ZINC13526726-1.pdb	-55.1828	-33.0629	-22.1199	0	21.1818
VYKSPNAYTLFS-ZINC13526723-1.pdb	-61.2299	-33.0689	-28.1609	0	20.5
VYKSPNAYTLFS-ZINC13526037-1.pdb	-50.6874	-36.1574	-14.53	0	22.6
VYKSPNAYTLFS-ZINC13488689-0.pdb	-59.3841	-56.8841	-2.5	0	18.8889
VYKSPNAYTLFS-ZINC13488688-1.pdb	-54.9134	-49.1458	-5.76754	0	18.5556
VYKSPNAYTLFS-ZINC13472694-1.pdb	-56.0021	-42.4199	-13.5822	0	18.4667
VYKSPNAYTLFS-ZINC13472690-1.pdb	-60.7869	-46.7869	-14	0	21.4
VYKSPNAYTLFS-ZINC13403607-0.pdb	-54.0221	-54.0221	0	0	20.0625
VYKSPNAYTLFS-ZINC13074615-0.pdb	-53.485	-38.6268	-14.8583	0	18.4706
VYKSPNAYTLFS-ZINC13074613-0.pdb	-57.6943	-41.4395	-16.2547	0	18.5882
VYKSPNAYTLFS-ZINC13074611-1.pdb	-53.5587	-41.0196	-12.539	0	18.5882
VYKSPNAYTLFS-ZINC13074609-0.pdb	-55.912	-45.5203	-10.3917	0	18.4706
VYKSPNAYTLFS-ZINC12956364-0.pdb	-52.1016	-38.1016	-14	0	20.2727
VYKSPNAYTLFS-ZINC12956361-0.pdb	-57.5513	-37.5969	-19.9544	0	19.7333
VYKSPNAYTLFS-ZINC12650514-0.pdb	-57.6621	-42.3013	-15.3608	0	19.0588
VYKSPNAYTLFS-ZINC12650510-1.pdb	-57.8428	-49.8811	-7.96176	0	18.5294
VYKSPNAYTLFS-ZINC12648245-1.pdb	-48.3462	-32.9536	-15.3927	0	21.2
VYKSPNAYTLFS-ZINC12474785-0.pdb	-53.3958	-41.841	-11.5548	0	18.4706
VYKSPNAYTLFS-ZINC12474718-1.pdb	-52.2604	-35.3545	-17.1613	0.255452	18.6429
VYKSPNAYTLFS-ZINC12471325-1.pdb	-53.8233	-44.2176	-9.60574	0	19.5
VYKSPNAYTLFS-ZINC12471319-1.pdb	-54.1531	-43.6531	-10.5	0	19.125
VYKSPNAYTLFS-ZINC12471311-0.pdb	-52.1073	-35.4499	-16.6574	0	18.5
VYKSPNAYTLFS-ZINC12471304-1.pdb	-53.1608	-37.7079	-15.4528	0	18.4375
VYKSPNAYTLFS-ZINC08983875-1.pdb	-53.9754	-46.9983	-6.97715	0	20.2
VYKSPNAYTLFS-ZINC08700018-0.pdb	-54.6343	-35.0994	-19.535	0	22.4545
VYKSPNAYTLFS-ZINC08700008-0.pdb	-54.7837	-44.2837	-10.5	0	21.2308
VYKSPNAYTLFS-ZINC08699996-1.pdb	-50.4152	-35.9136	-14.5015	0	22.2727
VYKSPNAYTLFS-ZINC08655204-1.pdb	-47.1112	-37.4294	-9.68182	0	21.3333
VYKSPNAYTLFS-ZINC08450977-0.pdb	-49.1015	-38.2059	-10.8956	0	20.8333
VYKSPNAYTLFS-ZINC08242850-1.pdb	-49.1896	-45.6896	-3.5	0	18.4706
VYKSPNAYTLFS-ZINC08034719-0.pdb	-39.9988	-26.9216	-13.0772	0	27.4286
VYKSPNAYTLFS-ZINC08034705-1.pdb	-57.2224	-42.625	-14.9697	0.372305	18.6429
VYKSPNAYTLFS-ZINC06925612-0.pdb	-59.9574	-27.5876	-32.3698	0	19
VYKSPNAYTLFS-ZINC06825014-0.pdb	-54.1229	-40.1229	-14	0	20.4167
VYKSPNAYTLFS-ZINC06824972-0.pdb	-53.7631	-37.2417	-16.5214	0	20
VYKSPNAYTLFS-ZINC06816987-1.pdb	-59.8163	-35.5681	-24.2482	0	21.5833
VYKSPNAYTLFS-ZINC06816982-0.pdb	-57.3202	-38.9634	-18.3568	0	22.3333

VYKSPNAYTLFS-ZINC06754980-0.pdb	-65.12	-36.7347	-29.0945	0.709139	18.7059
VYKSPNAYTLFS-ZINC06753324-1.pdb	-59.9482	-47.9482	-12	0	19.2
VYKSPNAYTLFS-ZINC06691151-1.pdb	-69.8395	-53.9597	-15.8797	0	21.4118
VYKSPNAYTLFS-ZINC06620297-0.pdb	-54.7395	-40.9409	-13.7987	0	18.5714
VYKSPNAYTLFS-ZINC06480800-0.pdb	-52.578	-34.6203	-18.322	0.364353	22.4545
VYKSPNAYTLFS-ZINC06472192-1.pdb	-55.9554	-41.9554	-14	0	19.0714
VYKSPNAYTLFS-ZINC06437595-1.pdb	-51.4972	-35.1074	-16.7668	0.377043	22.2727
VYKSPNAYTLFS-ZINC06379516-1.pdb	-46.5961	-37.1296	-9.46644	0	24.3333
VYKSPNAYTLFS-ZINC06361567-0.pdb	-54.5792	-23.0518	-31.5274	0	18.5714
VYKSPNAYTLFS-ZINC06361557-0.pdb	-54.5289	-30.6088	-23.9201	0	18.5714
VYKSPNAYTLFS-ZINC06361538-1.pdb	-53.0359	-42.0791	-10.9568	0	18.5714
VYKSPNAYTLFS-ZINC06361530-0.pdb	-51.8224	-21.3132	-30.5091	0	18.5
VYKSPNAYTLFS-ZINC06350042-0.pdb	-48.0007	-37.7528	-10.2479	0	19.3077
VYKSPNAYTLFS-ZINC06295961-1.pdb	-61.2949	-41.791	-19.504	0	19.0588
VYKSPNAYTLFS-ZINC06295960-0.pdb	-62.6464	-37.4616	-25.1849	0	18.6471
VYKSPNAYTLFS-ZINC06295954-1.pdb	-57.2337	-35.4463	-21.7873	0	18.5294
VYKSPNAYTLFS-ZINC06295953-1.pdb	-56.1544	-41.904	-14.2504	0	18.5882
VYKSPNAYTLFS-ZINC06180333-0.pdb	-56.4463	-44.8925	-11.5538	0	20.1333
VYKSPNAYTLFS-ZINC06175402-1.pdb	-51.666	-34.082	-17.5839	0	21.7692
VYKSPNAYTLFS-ZINC06117090-0.pdb	-53.2591	-46.2591	-7	0	19.1333
VYKSPNAYTLFS-ZINC06117083-1.pdb	-55.1897	-44.3887	-10.8011	0	18.4706
VYKSPNAYTLFS-ZINC06093529-0.pdb	-61.4247	-42.6295	-18.7952	0	22.3846

Table 2:

VYKSPNAYTLFS-ZINC17063815-0.pdb	-60.1013	-43.2365	-16.8648	0	21.0714
VYKSPNAYTLFS-ZINC17047258-1.pdb	-58.3489	-43.3784	-14.9705	0	19.6429
VYKSPNAYTLFS-ZINC17001182-1.pdb	-55.4565	-40.9074	-14.5491	0	21.8462
VYKSPNAYTLFS-ZINC16997856-1.pdb	-49.1275	-32.4968	-16.9452	0.314591	20.4545
VYKSPNAYTLFS-ZINC16982373-0.pdb	-65.1802	-56.8691	-8.31107	0	21.3889
VYKSPNAYTLFS-ZINC16968896-0.pdb	-55.1697	-50.0923	-5.07744	0	20.7857
VYKSPNAYTLFS-ZINC16955443-1.pdb	-55.8506	-41.8508	-13.9998	0	20.8462
VYKSPNAYTLFS-ZINC16954002-1.pdb	-43.4843	-39.9886	-3.49564	0	19.5385
VYKSPNAYTLFS-ZINC16952697-0.pdb	-46.2516	-42.7516	-3.5	0	18.4706
VYKSPNAYTLFS-ZINC16952695-1.pdb	-49.7854	-46.3098	-3.47561	0	19.5333
VYKSPNAYTLFS-ZINC16941446-0.pdb	-53.9915	-40.9138	-13.0777	0	19.3125
VYKSPNAYTLFS-ZINC16941442-1.pdb	-51.6702	-38.1111	-13.5591	0	21.3077
VYKSPNAYTLFS-ZINC16940978-1.pdb	-44.722	-25.1469	-19.5751	0	20.625
VYKSPNAYTLFS-ZINC16940076-1.pdb	-56.3418	-41.3814	-14.9604	0	20.6667
VYKSPNAYTLFS-ZINC16923303-0.pdb	-51.0758	-48.5812	-2.49463	0	18.8125
VYKSPNAYTLFS-ZINC16696935-1.pdb	-56.3467	-46.8853	-9.46141	0	20.0588
VYKSPNAYTLFS-ZINC16696932-0.pdb	-55.2488	-44.152	-11.0968	0	20.75
VYKSPNAYTLFS-ZINC16696930-0.pdb	-62.1698	-45.3272	-16.8426	0	19.5882

VYKSPNAYTLFS-ZINC16123805-1.pdb	-51.2624	-38.6561	-12.6064	0	20.9167
VYKSPNAYTLFS-ZINC15937364-0.pdb	-56.0892	-46.5892	-9.5	0	18.5
VYKSPNAYTLFS-ZINC15937228-1.pdb	-52.7566	-45.8227	-6.93396	0	18.7143
VYKSPNAYTLFS-ZINC15937227-0.pdb	-52.3191	-36.9014	-15.6965	0.27889	18.5
VYKSPNAYTLFS-ZINC15935031-1.pdb	-54.155	-42.8319	-11.3231	0	18.5625
VYKSPNAYTLFS-ZINC15922260-1.pdb	-55.2799	-45.8123	-9.46764	0	18.5294
VYKSPNAYTLFS-ZINC15894679-1.pdb	-61.4056	-43.9606	-17.445	0	20.8
VYKSPNAYTLFS-ZINC15785332-1.pdb	-52.3768	-45.3768	-7	0	18.5333
VYKSPNAYTLFS-ZINC15785118-0.pdb	-52.4424	-37.7134	-14.7289	0	21.5455
VYKSPNAYTLFS-ZINC15784788-1.pdb	-48.6639	-36.252	-12.4119	0	19.6923
VYKSPNAYTLFS-ZINC15784786-0.pdb	-44.626	-33.453	-11.173	0	22.5
VYKSPNAYTLFS-ZINC15784784-1.pdb	-46.8674	-40.776	-6.09139	0	23.2727
VYKSPNAYTLFS-ZINC15784782-1.pdb	-46.434	-40.5153	-5.91874	0	23.3636
VYKSPNAYTLFS-ZINC15784780-1.pdb	-47.2555	-41.2121	-6.04336	0	23.3636
VYKSPNAYTLFS-ZINC15784546-0.pdb	-58.2402	-52.5056	-5.73467	0	20.4375
VYKSPNAYTLFS-ZINC15784544-0.pdb	-55.6665	-46.2791	-9.38742	0	21.0714
VYKSPNAYTLFS-ZINC15784358-0.pdb	-51.9988	-34.3997	-17.5991	0	20.8
VYKSPNAYTLFS-ZINC15784356-1.pdb	-50.8377	-47.3377	-3.5	0	20.9231
VYKSPNAYTLFS-ZINC15784354-1.pdb	-49.6606	-43.421	-6.2396	0	22.6667
VYKSPNAYTLFS-ZINC15783176-1.pdb	-55.645	-48.7516	-6.89332	0	20.625
VYKSPNAYTLFS-ZINC15782452-0.pdb	-53.6382	-34.0165	-19.6217	0	22.4545
VYKSPNAYTLFS-ZINC15781858-1.pdb	-64.1145	-50.1472	-13.9673	0	22.1429
VYKSPNAYTLFS-ZINC15781121-0.pdb	-55.4475	-41.9497	-13.4978	0	20.4615
VYKSPNAYTLFS-ZINC15780430-1.pdb	-66.5006	-40.3675	-26.1331	0	18.6
VYKSPNAYTLFS-ZINC15780220-0.pdb	-64.5033	-53.3748	-11.1285	0	19.1765
VYKSPNAYTLFS-ZINC15780218-0.pdb	-56.7917	-47.4859	-9.3058	0	20.5
VYKSPNAYTLFS-ZINC15779471-1.pdb	-56.3374	-42.3478	-13.9896	0	20.4167
VYKSPNAYTLFS-ZINC15778754-1.pdb	-56.5915	-46.0934	-10.4981	0	19.6
VYKSPNAYTLFS-ZINC15778697-0.pdb	-63.4126	-45.515	-18.1577	0.260111	19.4
VYKSPNAYTLFS-ZINC15443135-0.pdb	-51.5021	-37.5021	-14	0	20.9091
VYKSPNAYTLFS-ZINC15443133-1.pdb	-48.8158	-34.8668	-13.949	0	21.3636
VYKSPNAYTLFS-ZINC15443126-0.pdb	-62.1473	-32.5383	-29.609	0	20.0909
VYKSPNAYTLFS-ZINC15440710-1.pdb	-47.9766	-40.9766	-7	0	19.6923
VYKSPNAYTLFS-ZINC15440708-1.pdb	-45.0148	-38.0352	-6.97963	0	19.0769
VYKSPNAYTLFS-ZINC15022795-0.pdb	-53.7064	-40.5723	-13.134	0	23.0769
VYKSPNAYTLFS-ZINC15022783-0.pdb	-54.0612	-43.571	-10.4902	0	18.6
VYKSPNAYTLFS-ZINC15022779-0.pdb	-48.6737	-46.1737	-2.5	0	21.5833
VYKSPNAYTLFS-ZINC14983149-1.pdb	-59.7013	-53.7194	-5.98191	0	19.7222

From Tables 1,2 it is easy to see that the number of iterations and the number of function evaluations of the new method with CG are less than those of the normal method for these test problems. Moreover, the cpu time and the final function norm evaluations of the new method with CG are more competitive than those of the normal method. For the VIM1 method, the results of Problems 1-7 are very interesting, but it fails for Problems 8-10. Moreover, it is not difficult to find that more CUP time is needed for this method. The main

reason maybe lies in the computation of the Jacobian matrix at every titeration.

Numerical results of VIM1 method

The tool of Dolan and Moré [63] is used to analyze the efficiency of these three algorithms.

Figures 1-3 show that the performances of these methods are relative to NI, NG, and cpu time of Tables 1,2 respectively. The numerical results indicate that the proposed method performs best among these three methods. To this end, we think that the enhancement of this proposed method is noticeable.

Performance profiles of these three methods (NG).

Function 5

Trigexp function [[57], p. 473]

f1(x)=3x13+2x2-5+sin(x1-x2)sin(x1+x2),fi(x)=-xi-1exi-1-xi+xi(4+3xi2)+2xi+1+sin(xi-xi+1)sin(xi+xi+1)-8,i=2,3,...,n-1,fn(x)=-xn-1exn-1-xn+4xn-3.

Initial guess: x0 = (0,0,...,0)T.

Function 6

Strictly convex function 1 [[52], p. 29]. e(x) is the gradient of h(x)=∑i=1n(exi-xi).

fi(x) = exi - 1, i = 1, 2, 3, ..., n.

Initial guess: x0= (1n,2n,...,1)T.

Function 7

Strictly convex function 2 [[52], p. 30]

e(x) is the gradient of h(x)=∑i=1ni10(exi-xi).

fi(x)=i10(exi-1),i=1,2,3,...,n.

Initial guess: x0 = (1,1,...,1)T.

Function 8

Variable dimensioned function

fi(x)=xi-1,i=1,2,3,...,n-2,fn-1(x)=∑j=1n-2j(xj-1),fn(x)=(∑j=1n-2j(xj-1))2.

Initial guess: x0=(1-1n,1-2n,...,0)T.

Function 9

Discrete boundary value problem [59]

f1(x)=2x1+0.5h2(x1+h)3-x2,fi(x)=2xi+0.5h2(xi+hi)3-xi-1+xi+1,i=2,3,...,n-1fn(x)=2xn+0.5h2(xn+hn)3-xn-1,h=1n+1.

Initial guess: x0 = (h(h - 1), h(2h - 1), ..., h(nh - 1)).

Function 10

The discretized two-point boundary value problem similar to the problem in [53]

e(x)=Ax+1(n+1)2F(x)=0,

when A is the n × n tridiagonal matrix given by

A=[8-1-18-1-18-1⋱⋱⋱⋱⋱-1-18],

and F(x) = (F1(x),F2(x),...,Fn(x))T with Fi(x) = sinxi - 1, i = 1, 2, ..., n, and x0 = (50, 0, 50, 0, ...).

In the experiments, all codes were written in MATLAB r2009a and run on a PC with G1620T@2.40 GHz CPU processor and 4.0 GB memory and Windows XP operation system. In order to compare the performance the given algorithm with CG's initial points (called new method with CG), we also do the experiment with only the main algorithm with initial points x0 (called the normal method). Aslam Noor et al. [60] presented a variational iteration technique for nonlinear equations, where the so-called VIM1 method has the better numerical performance. The VIM1 method has the following iteration form.

$$\sum_{k=0}^{\infty} \epsilon_k < \infty .$$

$$p(x_k + \alpha_k d_k) - p(x_k) \le \beta \alpha_k^2 e(x_k)^T d_k .$$

$$B_{k+1} = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k y_k^T}{y_k^T s_k} .$$

$$\begin{aligned} H_{k+1} &= H_k - \frac{y_k^T (s_k - H_k y_k) s_k s_k^T}{(y_k^T s_k)^2} + \frac{(s_k - H_k y_k) s_k^T + s_k (s_k - H_k y_k)^T}{(y_k^T s_k)^2} \\ &= \left(I - \frac{s_k y_k^T}{y_k^T s_k}\right) H_k \left(I - \frac{y_k s_k^T}{y_k^T s_k}\right) + \frac{s_k s_k^T}{y_k^T s_k} . \end{aligned}$$

$$e_k^T d_k < 0$$

$$\begin{aligned} \nabla p(x_k)^T d_k &= e(x_k)^T \nabla e(x_k) d_k \\ &= e(x_k)^T [(\nabla e(x_k) - B_k) d_k - e(x_k)] \\ &= e(x_k)^T (\nabla e(x_k) - B_k) d_k - e(x_k)^T e(x_k) . \end{aligned}$$

$$\begin{aligned} \nabla p(x_k)^T d_k + \| e(x_k) \|^2 &= e(x_k)^T (\nabla e(x_k) - B_k) d_k \\ &\le \| e(x_k) \| \| (\nabla e(x_k) - B_k) d_k \| . \end{aligned}$$

$$\begin{aligned} \nabla p(x_k)^T d_k &\le \| e(x_k) \| \| (\nabla e(x_k) - B_k) d_k \| - \| e(x_k) \|^2 \\ &\le -(1 - \epsilon_a) \| e(x_k) \|^2 . \end{aligned}$$

$$\lim_{k \to \infty} \| e_k \| = 0 .$$

$$p(x_{k+1}) - p(x_{l(k)}) \le \sigma \alpha_k e_k^T d_k < 0 .$$

$$b_2 \| d_k \|^2 \le d_k^T B_k d_k = -e_k^T d_k \le b_1 \| d_k \|^2 .$$

$$\begin{aligned} p(x_{l(k)}) &= p(x_{l(k)-1} + \alpha_{l(k)-1} d_{l(k)-1}) \\ &\le \max_{0 \le j \le m(l(k)-1)} \{ p(x_{l(k)-j-1}) \} + \sigma \alpha_{l(k)-1} g_{l(k)-1}^T d_{l(k)-1} \\ &\le \max_{0 \le j \le m(l(k)-1)} \{ p(x_{l(k)-j-1}) \} - \sigma b_2 \alpha_{l(k)-1} \| d_{l(k)-1} \|^2 . \end{aligned}$$

$$\begin{aligned} p(x_{l(k)}) &= p(x_{l(k)-1} + \alpha_{l(k)-1} d_{l(k)-1}) \\ &\le \max_{0 \le j \le m(l(k)-1)} \{ p(x_{l(k)-j-1}) \} + \sigma \alpha_{l(k)-1} g_{l(k)-1}^T d_{l(k)-1} \\ &\le \max_{0 \le j \le m(l(k)-1)} \{ p(x_{l(k)-j-1}) \} - \sigma b_2 \alpha_{l(k)-1} \| d_{l(k)-1} \|^2 . \end{aligned}$$

$$\lim_{k \to \infty} \alpha_{l(k)-1} \| d_{l(k)-1} \|^2 = 0 .$$

$$\lim_{k \to \infty} \inf d_{l(k)-1} = 0$$

$$\lim_{k \to \infty} \inf \alpha_{l(k)-1} = 0 .$$

$$\lim_{k \to \infty} \| d_{l(k)-j} \| = 0$$

$$\lim_{k \to \infty} p(x_{l(k)-j}) = \lim_{k \to \infty} p(x_{l(k)})$$

$$\lim_{k \to \infty} p(x_{l(k)}) = \lim_{k \to \infty} p(x_k) .$$

$$p(x_{k+1}) - p(x_{l(k)}) \le \alpha_k \sigma e_k^T d_k \le \alpha_k \sigma b_2 \| d_k \|^2 .$$

$$\lim_{k \to \infty} \alpha_k \| d_k \|^2 = 0,$$

$$\lim_{k \to \infty} \alpha_k = 0$$

$$\lim_{k \to \infty} \| d_k \| = 0 .$$

$$\alpha_k' = \frac{\alpha_k}{r}$$

$$p(x_k + \alpha_k' d_k) - p(x_k) \ge p(x_k + \alpha_k' d_k) - p(x_{l(k)}) > \sigma \alpha_k' e_k^T d_k .$$

$$p(x_k + \alpha_k' d_k) - p(x_k) = \alpha_k' \nabla p(x_k)^T d_k + o(\alpha_k' \| d_k \|) .$$

$$\nabla p(x_k)^T d_k = e_k^T \nabla e(x_k) d_k \le \delta^* e_k^T d_k .$$

$$[\delta^* - \sigma] \alpha_k' e_k^T d_k + o(\alpha_k' \| d_k \|) \ge 0 .$$

$$e_k^T d_k < 0$$

$$\alpha_k' \| d_k \|$$

$$\lim_{k \to \infty} \frac{e_k^T d_k}{\| d_k \|} = 0 .$$

$$\lim_{k\to\infty} \| d_k \| = 0.$$

$$a \equiv \left\{ \| \nabla e(x^*) \| + \tfrac{1}{2c}, 2c \right\}, c = \left\| \nabla e(x^*)^{-1} \right\|.$$

$$\tfrac{1}{a} \|x_k - x^*\| \le \|e(x_k)\| \le a\|x_k - x^*\|$$

$$B_k(x_{k+1} - x^*) + [\nabla e(x_k)(x_k - x^*) - B_k(x_k - x^*)]$$

$$+ [e(x_k) - e(x^*) - \nabla e(x_k)(x_k - x^*)]$$

$$= e(x_k) + B_k d_k = 0,$$

$$f_1(x) = e^{x_1} - 1,$$

$$f_i(x) = \tfrac{i}{10}(e^{x_i} + x_{i-1} - 1), i = 2, 3, \ldots, n.$$

$$x_0 = \left(\tfrac{1}{n^2}, \tfrac{1}{n^2}, \ldots, \tfrac{1}{n^2}\right)^T$$

$$f_i(x) = 2\left(n + i(1 - \cos x_i) - \sin x_i - \sum_{j=1}^{n} \cos x_j\right)(2 \sin x_i - \cos x_i), i$$

$$= 1, 2, 3, \ldots, n.$$

$$x_0 = \left(\tfrac{101}{100n}, \tfrac{101}{100n}, \ldots, \tfrac{101}{100n}\right)^T$$

$$f_i(x) = \ln(x_i + 1) - \tfrac{x_i}{n}, i = 1, 2, 3, \ldots, n.$$

$$f_1(x) = (3 - 0.5x_1)x_1 - 2x_2 + 1,$$

$$f_i(x) = (3 - 0.5x_i)x_i - x_{i-1} + 2x_{i+1} + 1,$$

$$i = 2, 3, \ldots, n - 1,$$

$$f_n(x) = (3 - 0.5x_n)x_n - x_{n-1} + 1.$$

$$f_1(x) = 3x_1^3 + 2x_2 - 5 + \sin(x_1 - x_2)\sin(x_1 + x_2),$$

$$f_i(x) = -x_{i-1}e^{x_{i-1} - x_i} + x_i(4 + 3x_i^2) + 2x_{i+1}$$

$$+ \sin(x_i - x_{i+1})\sin(x_i + x_{i+1}) - 8, i = 2, 3, \ldots, n - 1,$$

$$h(x) = \sum_{i=1}^{n} (e^{x_i} - x_i)$$

$$x_0 = \left(\tfrac{1}{n}, \tfrac{2}{n}, \ldots, 1\right)^T$$

$$h(x) = \sum_{i=1}^{n} \tfrac{i}{10}(e^{x_i} - x_i)$$

$$f_i(x) = \tfrac{i}{10}(e^{x_i} - 1), i = 1, 2, 3, \ldots, n.$$

$$f_i(x) = x_i - 1, i = 1, 2, 3, \ldots, n - 2,$$

$$f_{n-1}(x) = \sum_{j=1}^{n-2} j(x_j - 1),$$

$$f_n(x) = \left(\sum_{j=1}^{n-2} j(x_j - 1)\right)^2.$$

$$x_0 = \left(1 - \tfrac{1}{n}, 1 - \tfrac{2}{n}, \ldots, 0\right)^T$$

$$f_1(x) = 2x_1 + 0.5h^2(x_1 + h)^3 - x_2,$$

$$f_i(x) = 2x_i + 0.5h^2(x_i + hi)^3 - x_{i-1} + x_{i+1},$$

$$i = 2, 3, \ldots, n - 1$$

$$f_n(x) = 2x_n + 0.5h^2(x_n + hn)^3 - x_{n-1},$$

$$h = \frac{1}{n+1}.$$

$$e(x) = Ax + \frac{1}{(n+1)^2} F(x) = 0,$$

$$A = \begin{bmatrix} 8 & -1 & & & & \\ -1 & 8 & -1 & & & \\ & -1 & 8 & -1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & \ddots & \ddots & -1 \\ & & & & -1 & 8 \end{bmatrix},$$

$$y_k^T s_k > 0$$

$$d_k^T e_k < 0$$

$$d_k^T \nabla e(x_k) e_k < 0$$

$$d_k = \begin{cases} -e_k + \dfrac{e_k^T (e_k - e_{k-1})}{\|e_{k-1}\|^2} d_{k-1} & \text{if } k \geq 1, \\ -e_k & \text{if } k = 0. \end{cases}$$

$$\epsilon_k = \frac{1}{Nl^2}$$

Results

Statistical Analysis and Validation

CoMFA and CoMSIA approaches were carried out using the 40 1,7-diazacarbazole derivatives chosen in the Chk1 inhibition experiment. The statistical summaries are shown in Table 2. In the CoMFA model, the PLS regression analysis yielded a q2 of 0.726 with SEE of 0.215. Then, the non-cross-validated method gave an r2 value of 0.918, an F-value of 115.292 and three optimal components (Figure 3).

In the CoMSIA model, a q2 value of 0.531 and r2 value of 0.950 were obtained when five field descriptors were considered [313]. The F-value, SEE value, and ONC values were 1413412, 0.1713, and 4, respectively. From the field contribution results, we found that the electrostatic field played a major contribution among the present fields. These data also showed the reliability of the CoMSIA model.

A test set of 40 compounds were used to validate the accuracy and predictability of the model. The values of r2pred, shown in the Table 2, were 0.878 for CoMFA and 0.846 for the CoMSIA model. These statistical indices indicated a good external predictive capacity of the models. The graphs showing the experimental and predicted pIC50 values for the total set used in the CoMFA and CoMSIA approaches are described in Figure 4. The good linear relationships illustrated that the bioactivities predicted by the derived models were in agreement with the experimental data, indicating that these models had satisfactory predictive capacity (Figure 4).

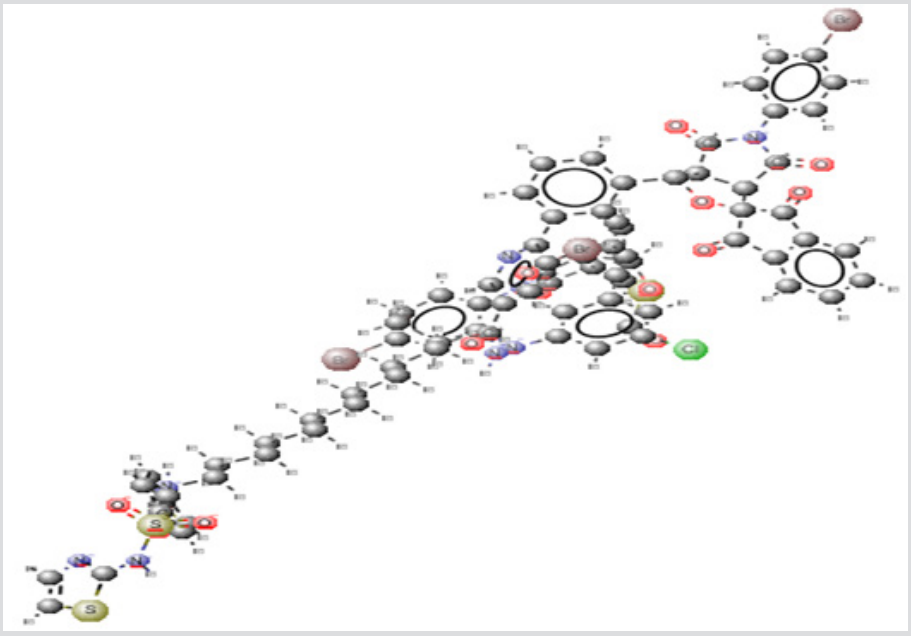


Figure 4: Plots of experimental vs. predicted pIC50 values for the total set in the CoMFA (A) and CoMSIA (B) models.

CoMFA/CoMSIA contour map analysis: The steady 3D QSAR models were generally applied to drug discovery processes to predict the biological activities of unknown derivatives. Moreover, the effects of the field descriptors contributing to activities can be partitioned and visualized through 3D contour plots. The field type of the contour plot was set to StDev × Coeff to aid visualization.

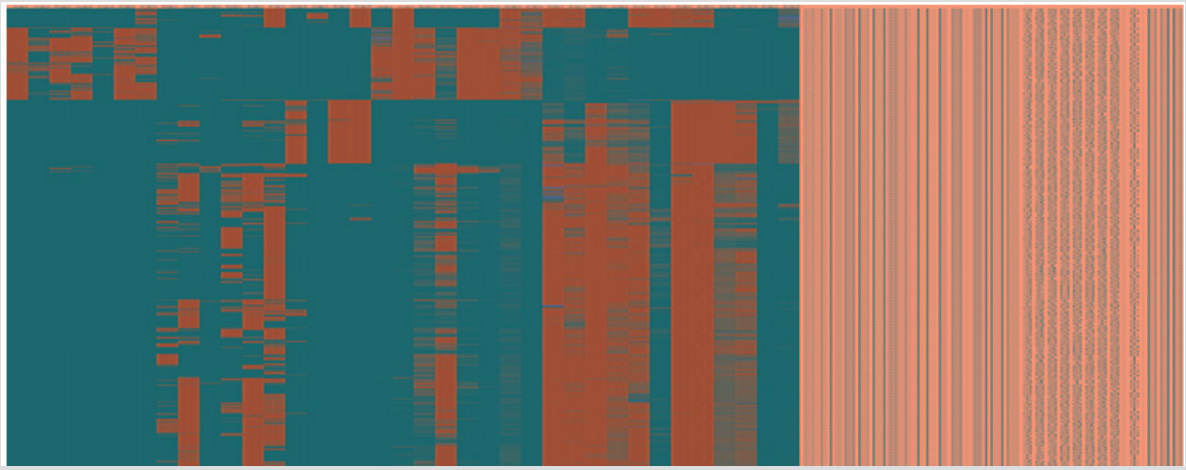


Figure 5

The CoMFA contour maps are illustrated in Figure 5. These contours plots demonstrate regions where the steric and electrostatic variations in the different molecular features lead to either increased or reduced bioactivity. The most structure 28 was chosen as a reference to aid visualization. Two green regions and two yellow places exist around the compound zones represented the steric favorable and unfavorable areas, respectively [314]. The green maps around the methyl group of pyrazole ring indicated that bulk groups were favored there. The methyl group of pyrazole ring at this position may be favorable to the interaction between the compound and its receptor. It can be explained by the fact that the bioactivity of compound 28 (pIC50 = 9.51) was better than those of compound 27 (pIC50 = 8.56). A yellow contour around the R1substituent indicated that the small bulky group can enhance bioactivities. This phenomenon could be concluded by comparing the activities between compounds 25 (pIC50 = 9.34) and 30 (pIC50 = 8.64) (Figure 5,6) (Table 1).

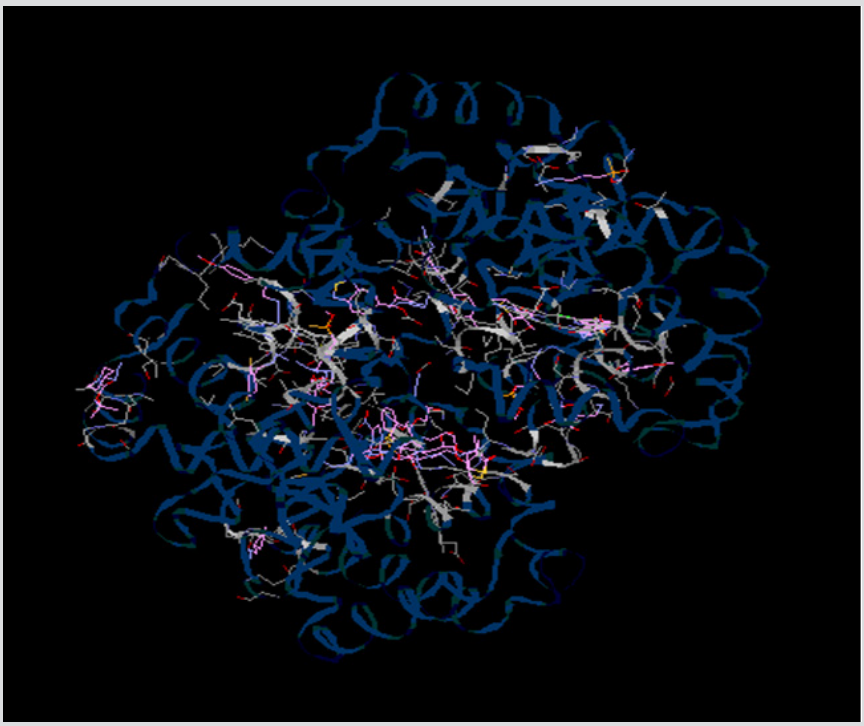


Figure 6: Steric (A) and electrostatic (B) contours of the 23,45,76, CoMFA model.

The green color shows the favored steric area and the yellow color show steric area. The red color shows the favored negative electrostatic area and the blue color shows the favored positive electrostatic area.

In the CoMFA electrostatic contour maps, the regions in red implied where electronegative groups improved activities, whereas the positions in blue purported where electropositive groups enhanced activities. A blue contour around the methyl group of pyrazole ring, indicating a positive atomic charge group in this position had a positive effect on the inhibitory activity, such as compounds 28 and 27. Moreover, a large blue tetrahedron around the para-position of the piperidine ring revealed the importance of positive atomic groups. For example, compounds 23,45,76, shown moderate activity probably because the para-position of piperidine ring was replaced by

electronegative groups. Simultaneously, two small red cubes near the nitrogen atoms of the pyrazole and piperidine rings indicated that the negatively-charged groups were helpful for inhibitory activity, which was consistent with the experimental results.

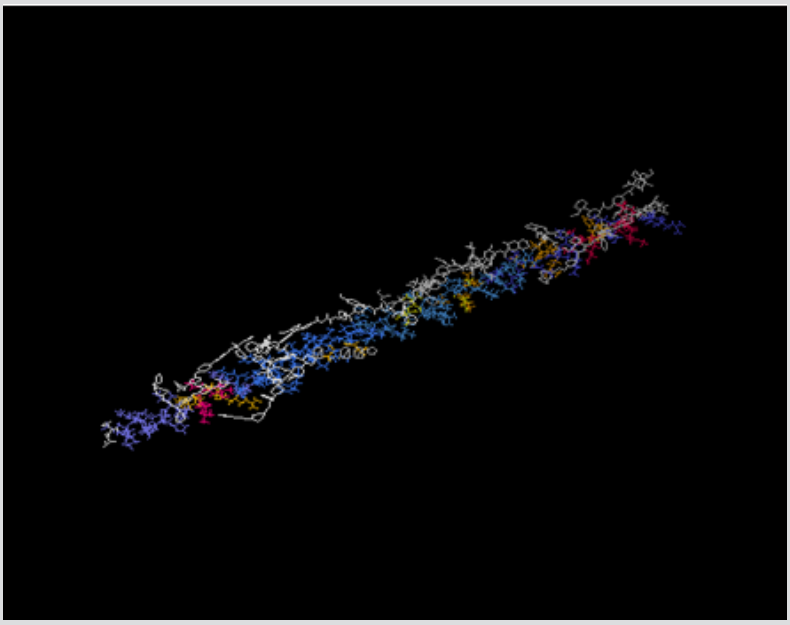


Figure 7: Steric (A) and electrostatic (B) contours of the CoMSIA 23,45,76, model.

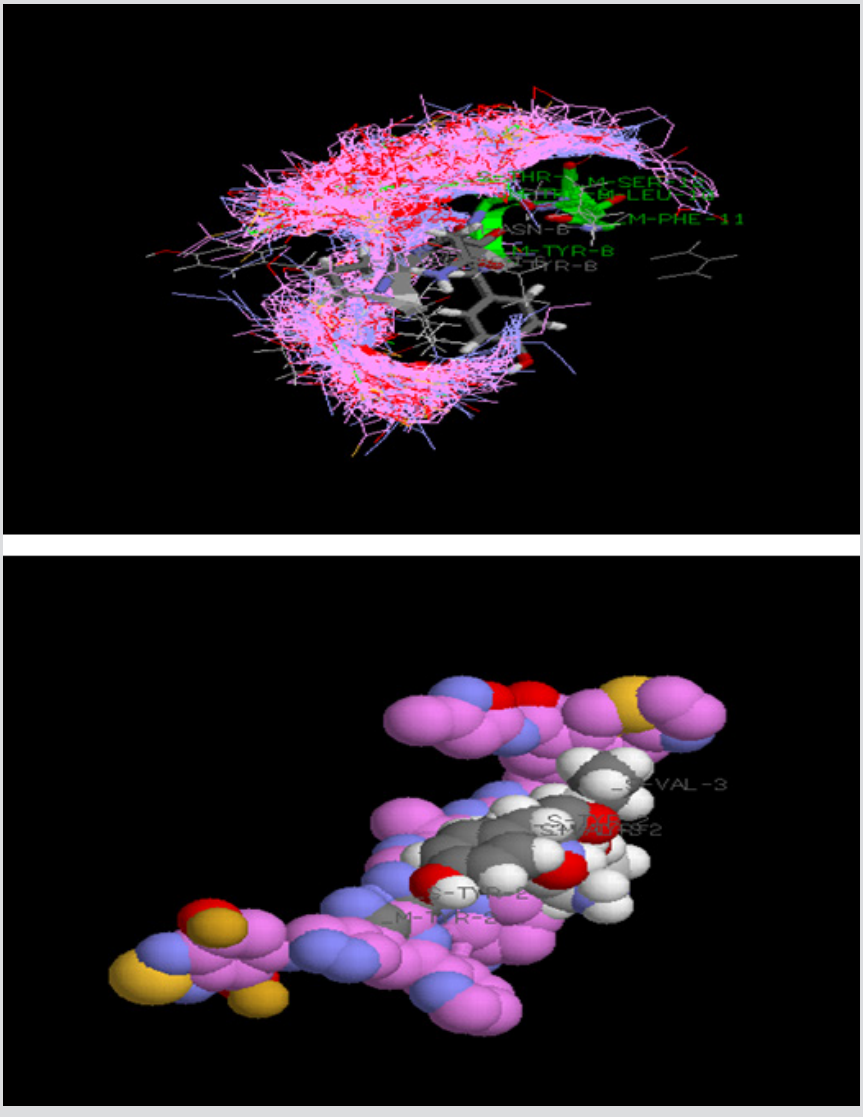


Figure 8: Hydrophobic contours of the CoMSIA 23,45,76, model. The yellow color shows the favored hydrophobic area, the white color shows the disfavored hydrophobic area.

The CoMSIA steric and electrostatic contour plots were described in Figure 7. Not surprisingly, most contours were similar to those of CoMFA model and, hence, were not discussed. The hydrophobic contour plot was constructed through CoMSIA model, presented in Figure 8. The yellow and gray colors represented favorable and unfavorable hydrophobic areas, respectively. The hydrophobic contour maps were exclusively located around R1 and R2 substituents. A small yellow cube surrounding the C-1 position of pyrazole ring (R1 group) indicated hydrophobic groups were advantageous for inhibitory activity. Compounds 25-29 showed better inhibitory activities. Reaction-diffusion Chemoproteomic-aided Molecular designed, ovarian stem cells (OSCs) motif VYKSPNAYTLFS derived and a quasi-Newton algorithm Scan Predicted antigenic peptide-like polypharmacophoric ligand (CAMSPCPL) targeted on the conserved (β2GPI) antibodies, Anti-phospholipid antibodies (aPL) or from (PAPP-A; gi:38045915) and APOH apolipoprotein H post-trancripts domains to Germline low energetic than compounds 9-15, probably because -CN group was replaced by hydrophobic substituents [315]. A yellow cone around the meta-position of benzene ring revealed that hydrophobic substituents in this region produced increased bioactivity. The gray contours around the para-position of the piperidine ring (R2 substituent) suggested that the hydrophilic group in this region could increase activity. It can be demonstrated by the fact that the biological activity of compound 11 (pIC50 = 8.745) was slightly higher than that of compound 13 (pIC50 = 8.301) (Table 2,3 and Figure 7,8).

Table 3:

VYKSPNAYTLFS-ZINC34498487-0.pdb	-59.1756	-48.6374	-10.5383	0	20.4118
VYKSPNAYTLFS-ZINC34498401-1.pdb	-56.3614	-41.821	-14.5404	0	19.9231
VYKSPNAYTLFS-ZINC34498399-0.pdb	-57.851	-34.7941	-23.0569	0	18.5625
VYKSPNAYTLFS-ZINC34496883-0.pdb	-47.8991	-45.3991	-2.5	0	18.4706
VYKSPNAYTLFS-ZINC34492113-0.pdb	-52.7199	-32.3432	-20.6584	0.281662	18.5
VYKSPNAYTLFS-ZINC34492109-0.pdb	-53.8359	-40.9908	-12.8452	0	20.5714
VYKSPNAYTLFS-ZINC34492045-1.pdb	-57.4097	-48.7343	-8.67533	0	21.6667
VYKSPNAYTLFS-ZINC34490273-1.pdb	-62.4344	-50.4344	-12	0	19.1176
VYKSPNAYTLFS-ZINC34490069-1.pdb	-59.9896	-52.9896	-7	0	20.1765
VYKSPNAYTLFS-ZINC34490067-1.pdb	-59.2848	-52.6248	-6.66007	0	18.5556
VYKSPNAYTLFS-ZINC34490065-0.pdb	-63.3981	-56.3981	-7	0	21.0588
VYKSPNAYTLFS-ZINC34489224-0.pdb	-45.3769	-32.2978	-13.4567	0.377568	23.2222
VYKSPNAYTLFS-ZINC34479753-1.pdb	-52.4461	-45.4461	-7	0	20
VYKSPNAYTLFS-ZINC34477542-0.pdb	-62.7036	-52.2036	-10.5	0	20
VYKSPNAYTLFS-ZINC34477541-1.pdb	-50.2019	-34.4061	-16.0702	0.274371	20.5
VYKSPNAYTLFS-ZINC34477540-0.pdb	-47.5876	-36.367	-11.2206	0	22.7273
VYKSPNAYTLFS-ZINC34475831-0.pdb	-58.7455	-49.2455	-9.5	0	18.75
VYKSPNAYTLFS-ZINC34475830-0.pdb	-62.7724	-52.7126	-10.0598	0	20.5
VYKSPNAYTLFS-ZINC34475829-1.pdb	-59.5661	-50.0661	-9.5	0	19.1875
VYKSPNAYTLFS-ZINC34475828-1.pdb	-59.6321	-51.1405	-8.49168	0	20.3125
VYKSPNAYTLFS-ZINC34469114-1.pdb	-65.3235	-48.059	-17.2646	0	19.5556
VYKSPNAYTLFS-ZINC34464194-1.pdb	-49.7706	-36.9625	-12.8081	0	22.8
VYKSPNAYTLFS-ZINC34464193-1.pdb	-51.9414	-35.6523	-16.2891	0	25
VYKSPNAYTLFS-ZINC34464192-0.pdb	-47.6577	-33.5617	-14.0959	0	24.1
VYKSPNAYTLFS-ZINC34464191-1.pdb	-56.608	-26.3173	-30.6302	0.339485	21
VYKSPNAYTLFS-ZINC34462809-0.pdb	-46.4501	-25.2449	-21.2053	0	19.3333
VYKSPNAYTLFS-ZINC34454786-0.pdb	-44.8621	-32.3533	-12.8635	0.354718	24.1111
VYKSPNAYTLFS-ZINC34454785-1.pdb	-42.2837	-35.2837	-7	0	22.3333

VYKSPNAYTLFS-ZINC34452039-1.pdb	-57.6843	-40.7577	-16.9266	0	22.6667
VYKSPNAYTLFS-ZINC34433807-1.pdb	-51.7121	-38.2664	-13.4457	0	20.625
VYKSPNAYTLFS-ZINC34427725-1.pdb	-59.6897	-45.8133	-13.8765	0	19.7143
VYKSPNAYTLFS-ZINC34417532-1.pdb	-69.3807	-52.6936	-16.6871	0	19.5556
VYKSPNAYTLFS-ZINC34417531-1.pdb	-70.0355	-47.8721	-22.1633	0	18.9444
VYKSPNAYTLFS-ZINC34412143-1.pdb	-56.8141	-51.3822	-5.43189	0	18.8235
VYKSPNAYTLFS-ZINC34406717-1.pdb	-52.9359	-49.4359	-3.5	0	19
VYKSPNAYTLFS-ZINC34370211-0.pdb	-53.2251	-40.1289	-13.0961	0	18.75
VYKSPNAYTLFS-ZINC34370210-0.pdb	-54.5979	-40.0206	-14.5773	0	18.5
VYKSPNAYTLFS-ZINC34370209-1.pdb	-57.4226	-38.5039	-18.9187	0	19.1875
VYKSPNAYTLFS-ZINC34370208-1.pdb	-56.5261	-41.7385	-14.7876	0	18.5
VYKSPNAYTLFS-ZINC34329749-0.pdb	-56.8711	-38.5208	-18.3503	0	23.6364
VYKSPNAYTLFS-ZINC34329748-1.pdb	-65.4229	-33.093	-32.3299	0	19.6364
VYKSPNAYTLFS-ZINC34321021-0.pdb	-55.644	-45.1921	-10.4519	0	19.3333
VYKSPNAYTLFS-ZINC34321020-0.pdb	-51.5124	-41.0188	-10.4936	0	18.7333
VYKSPNAYTLFS-ZINC34321019-1.pdb	-57.0079	-46.6158	-10.3921	0	19.0625
VYKSPNAYTLFS-ZINC34321018-0.pdb	-51.8658	-41.3658	-10.5	0	18.5
VYKSPNAYTLFS-ZINC34304487-1.pdb	-51.5188	-33.0929	-18.4259	0	20.8182
VYKSPNAYTLFS-ZINC34288034-1.pdb	-42.167	-32.667	-9.5	0	25.1429
VYKSPNAYTLFS-ZINC34287807-0.pdb	-59.7291	-43.7037	-16.0253	0	18.7059
VYKSPNAYTLFS-ZINC34287806-0.pdb	-59.2658	-44.8873	-14.3786	0	19
VYKSPNAYTLFS-ZINC34287805-1.pdb	-59.7547	-37.3127	-22.6814	0.239372	18.4706
VYKSPNAYTLFS-ZINC34287804-0.pdb	-50.4024	-33.2785	-17.1239	0	18.4706
VYKSPNAYTLFS-ZINC34278228-0.pdb	-38.7919	-31.8072	-6.9848	0	25.1429
VYKSPNAYTLFS-ZINC34235675-0.pdb	-56.1387	-43.9834	-12.1552	0	21.7333
VYKSPNAYTLFS-ZINC34182444-0.pdb	-52.1436	-39.7402	-12.4034	0	18.8824
VYKSPNAYTLFS-ZINC34161425-1.pdb	-39.9679	-25.9679	-14	0	21.8571
VYKSPNAYTLFS-ZINC34160134-0.pdb	-45.693	-30.6738	-15.3941	0.374855	21.4444
VYKSPNAYTLFS-ZINC34160133-0.pdb	-45.3508	-31.6934	-14.0211	0.363644	20.5556
VYKSPNAYTLFS-ZINC34139950-1.pdb	-52.4027	-28.9016	-23.5011	0	23.2222
VYKSPNAYTLFS-ZINC34139949-1.pdb	-47.8277	-36.1509	-11.6768	0	23.5556
VYKSPNAYTLFS-ZINC34139948-1.pdb	-44.6049	-27.4852	-17.1197	0	21.4444
VYKSPNAYTLFS-ZINC34129110-1.pdb	-56.7498	-46.2532	-10.4967	0	21.2857
VYKSPNAYTLFS-ZINC34122494-0.pdb	-56.3628	-46.1463	-10.2165	0	19.5333
VYKSPNAYTLFS-ZINC34116219-0.pdb	-56.4088	-41.8374	-14.5713	0	18.5714
VYKSPNAYTLFS-ZINC34082726-0.pdb	-53.9616	-34.9616	-19	0	18.5882
VYKSPNAYTLFS-ZINC34076870-1.pdb	-64.7002	-51.7135	-12.9867	0	21.5333

VYKSPNAYTLFS-ZINC34076869-0.pdb	-57.5791	-48.0983	-9.48076	0	19.1333
VYKSPNAYTLFS-ZINC34075807-1.pdb	-56.5795	-33.2224	-23.3572	0	23
VYKSPNAYTLFS-ZINC34027303-0.pdb	-47.8933	-34.3132	-13.58	0	25.4444
VYKSPNAYTLFS-ZINC33947808-0.pdb	-59.9722	-49.9374	-10.0348	0	20.6667
VYKSPNAYTLFS-ZINC33939040-1.pdb	-60.3517	-34.6725	-25.6791	0	20.2727
VYKSPNAYTLFS-ZINC33765559-1.pdb	-63.7726	-51.7541	-12.0185	0	19.3333
VYKSPNAYTLFS-ZINC33739862-0.pdb	-46.9141	-43.4141	-3.5	0	20.5
VYKSPNAYTLFS-ZINC33693876-0.pdb	-49.0912	-42.1699	-6.92134	0	22.75
VYKSPNAYTLFS-ZINC33506585-1.pdb	-56.2692	-36.0123	-20.2568	0	20.4286
VYKSPNAYTLFS-ZINC33378647-0.pdb	-61.9392	-53.1482	-8.79091	0	20.8235
VYKSPNAYTLFS-ZINC33376898-1.pdb	-54.2697	-37.8273	-16.4423	0	22.8571
VYKSPNAYTLFS-ZINC33376445-1.pdb	-50.8112	-38.349	-12.4622	0	19.0833
VYKSPNAYTLFS-ZINC32689261-1.pdb	-50.6494	-43.6947	-6.95477	0	19.4667
VYKSPNAYTLFS-ZINC32603235-0.pdb	-51.6817	-32.7103	-18.9714	0	22.5556
VYKSPNAYTLFS-ZINC32222731-0.pdb	-49.6471	-43.4595	-6.18753	0	22.8333
VYKSPNAYTLFS-ZINC32181558-0.pdb	-35.0443	-28.0443	-7	0	25.3333
VYKSPNAYTLFS-ZINC32180809-0.pdb	-56.4687	-46.1673	-10.3014	0	20.1333
VYKSPNAYTLFS-ZINC32107279-1.pdb	-61.507	-45.5513	-15.9558	0	18.4706
VYKSPNAYTLFS-ZINC32107277-0.pdb	-64.8584	-50.8584	-14	0	19.1176
VYKSPNAYTLFS-ZINC32107226-0.pdb	-65.2239	-46.8035	-18.4204	0	21.4375
VYKSPNAYTLFS-ZINC32107224-0.pdb	-57.5354	-45.7228	-11.8126	0	19.0625
VYKSPNAYTLFS-ZINC32107130-0.pdb	-56.5202	-33.402	-23.1182	0	18.5
VYKSPNAYTLFS-ZINC31776901-1.pdb	-58.0356	-44.2085	-13.8271	0	23
VYKSPNAYTLFS-ZINC31517997-1.pdb	-61.7249	-47.7249	-14	0	18.5625
VYKSPNAYTLFS-ZINC29553629-0.pdb	-60.7952	-47.9351	-12.86	0	19
VYKSPNAYTLFS-ZINC29553625-0.pdb	-54.2368	-45.2793	-9.32326	0.365662	18.6471
VYKSPNAYTLFS-ZINC29546418-0.pdb	-39.8244	-32.8244	-7	0	22.5556
VYKSPNAYTLFS-ZINC29546415-1.pdb	-39.8699	-32.8699	-7	0	21.4444
VYKSPNAYTLFS-ZINC28826366-0.pdb	-58.9713	-56.9443	-2.02702	0	20.25
VYKSPNAYTLFS-ZINC28826350-1.pdb	-65.8386	-61.343	-4.49562	0	20.5882
VYKSPNAYTLFS-ZINC28826348-0.pdb	-58.3626	-49.2583	-9.35709	0.252819	18.7333
VYKSPNAYTLFS-ZINC28826345-1.pdb	-63.5876	-50.5955	-12.9922	0	19.2353
VYKSPNAYTLFS-ZINC28826344-0.pdb	-59.4598	-45.5682	-13.8916	0	20
VYKSPNAYTLFS-ZINC28826341-1.pdb	-56.6987	-47.8375	-8.86115	0	18.4706
VYKSPNAYTLFS-ZINC28826340-1.pdb	-55.7956	-52.4112	-3.38436	0	19.8667
VYKSPNAYTLFS-ZINC28759422-1.pdb	-53.1489	-45.4777	-7.67116	0	18.5

VYKSPNAYTLFS-ZINC28704829-1.pdb	-51.9942	-40.6594	-11.3348	0	20.5
VYKSPNAYTLFS-ZINC28100854-1.pdb	-61.2097	-34.51	-26.6997	0	18.625
VYKSPNAYTLFS-ZINC28100848-1.pdb	-63.1965	-45.6965	-17.5	0	18.5625
VYKSPNAYTLFS-ZINC28100843-1.pdb	-50.0768	-47.5768	-2.5	0	18.6875
VYKSPNAYTLFS-ZINC28100838-0.pdb	-50.1862	-37.3707	-12.8155	0	18.625
VYKSPNAYTLFS-ZINC27553251-1.pdb	-67.3421	-50.4213	-16.9208	0	19.9412
VYKSPNAYTLFS-ZINC27553248-0.pdb	-64.9099	-51.5704	-13.3395	0	19.4118
VYKSPNAYTLFS-ZINC27192423-0.pdb	-60.9165	-43.3704	-17.5461	0	18.75
VYKSPNAYTLFS-ZINC27187558-0.pdb	-57.1815	-42.0153	-15.1662	0	20
VYKSPNAYTLFS-ZINC27187553-0.pdb	-56.5711	-32.2236	-24.3475	0	18.5
VYKSPNAYTLFS-ZINC27185422-1.pdb	-60.0396	-45.0292	-15.0104	0	19.7333
VYKSPNAYTLFS-ZINC27185410-1.pdb	-55.0055	-36.8101	-18.1954	0	18.4667
VYKSPNAYTLFS-ZINC27076149-1.pdb	-61.314	-42.8193	-18.4947	0	18.5714
VYKSPNAYTLFS-ZINC26838668-1.pdb	-57.1133	-40.6868	-16.4266	0	20.25
VYKSPNAYTLFS-ZINC26838664-1.pdb	-59.997	-27.8735	-32.1235	0	19.75
VYKSPNAYTLFS-ZINC26829224-0.pdb	-59.6076	-41.8574	-17.7503	0	19.0667
VYKSPNAYTLFS-ZINC26829220-1.pdb	-59.3786	-38.4251	-20.9535	0	19.3333
VYKSPNAYTLFS-ZINC26672590-1.pdb	-65.8979	-51.0704	-14.8275	0	19.1111
VYKSPNAYTLFS-ZINC26672586-1.pdb	-67.6052	-48.2232	-19.382	0	19.5556
VYKSPNAYTLFS-ZINC26547916-0.pdb	-44.0852	-31.5297	-12.9111	0.355501	24.3333
VYKSPNAYTLFS-ZINC26547913-0.pdb	-45.3508	-31.3265	-14.4592	0.434894	19
VYKSPNAYTLFS-ZINC26547891-1.pdb	-62.4925	-42.5481	-19.9444	0	19.2857
VYKSPNAYTLFS-ZINC26547888-0.pdb	-60.1436	-35.6061	-24.5375	0	18.5333
VYKSPNAYTLFS-ZINC26546949-0.pdb	-47.7053	-40.7053	-7	0	18.4286
VYKSPNAYTLFS-ZINC26546943-1.pdb	-45.6143	-38.6143	-7	0	19
VYKSPNAYTLFS-ZINC26021496-1.pdb	-56.8729	-31.3557	-25.5171	0	19.3125
VYKSPNAYTLFS-ZINC26019861-0.pdb	-63.4081	-34.5044	-28.9038	0	19.2
VYKSPNAYTLFS-ZINC26012015-1.pdb	-60.2255	-46.3297	-13.8958	0	19.7059
VYKSPNAYTLFS-ZINC26012012-1.pdb	-61.4981	-39.6111	-21.887	0	18.5882
VYKSPNAYTLFS-ZINC26012005-1.pdb	-60.1715	-41.8705	-18.301	0	19.0588
VYKSPNAYTLFS-ZINC26011999-0.pdb	-59.9766	-38.5079	-21.4687	0	20.2353
VYKSPNAYTLFS-ZINC25721242-0.pdb	-57.4088	-41.9761	-15.4327	0	20.9286
VYKSPNAYTLFS-ZINC25627094-1.pdb	-59.6005	-48.0653	-11.5351	0	19.4706
VYKSPNAYTLFS-ZINC25626937-1.pdb	-64.2475	-47.768	-16.4795	0	18.7059
VYKSPNAYTLFS-ZINC22204334-0.pdb	-67.9142	-46.539	-21.3752	0	18.4706
VYKSPNAYTLFS-ZINC22204315-1.pdb	-65.1417	-46.3464	-18.7953	0	19.5294

VYKSPNAYTLFS-ZINC22203891-1.pdb	-62.9235	-54.2623	-8.66127	0	19.8235
VYKSPNAYTLFS-ZINC22198994-0.pdb	-56.8364	-28.0401	-28.7964	0	18.7143
VYKSPNAYTLFS-ZINC22130459-1.pdb	-66.4794	-50.8036	-15.6757	0	20.0556
VYKSPNAYTLFS-ZINC22130406-0.pdb	-58.7567	-47.6271	-11.1296	0	21.1333
VYKSPNAYTLFS-ZINC22114256-1.pdb	-51.1572	-43.0803	-8.07692	0	20.6154
VYKSPNAYTLFS-ZINC22064434-0.pdb	-59.515	-45.0644	-14.4507	0	18.8
VYKSPNAYTLFS-ZINC22064432-1.pdb	-60.3454	-46.3467	-13.9987	0	19.2
VYKSPNAYTLFS-ZINC22058744-1.pdb	-59.9075	-37.4412	-22.4662	0	18.5882
VYKSPNAYTLFS-ZINC22058740-0.pdb	-58.6471	-51.16	-7.48714	0	19.4118
VYKSPNAYTLFS-ZINC22058732-1.pdb	-60.4677	-44.6825	-15.7852	0	18.4667
VYKSPNAYTLFS-ZINC22056810-0.pdb	-52.3701	-34.7854	-17.5847	0	22.3636
VYKSPNAYTLFS-ZINC22049550-1.pdb	-54.8932	-40.3007	-14.5925	0	23.7273
VYKSPNAYTLFS-ZINC22049545-1.pdb	-52.2278	-38.1098	-14.118	0	20.4167
VYKSPNAYTLFS-ZINC22049482-1.pdb	-57.1642	-37.686	-19.4782	0	18.4706
VYKSPNAYTLFS-ZINC22049477-1.pdb	-57.2593	-37.3717	-19.8876	0	18.4706
VYKSPNAYTLFS-ZINC22049472-1.pdb	-59.5881	-45.9197	-13.6684	0	19.4706
VYKSPNAYTLFS-ZINC22049467-0.pdb	-57.5837	-39.6205	-17.9633	0	18.5882
VYKSPNAYTLFS-ZINC22049463-0.pdb	-62.0983	-44.7595	-17.3388	0	18.4706
VYKSPNAYTLFS-ZINC22049458-1.pdb	-57.8672	-44.4136	-13.4536	0	18.4706
VYKSPNAYTLFS-ZINC22049453-0.pdb	-55.8994	-40.9842	-14.9152	0	19.9412
VYKSPNAYTLFS-ZINC22049448-0.pdb	-57.842	-43.0461	-14.7959	0	20.2857
VYKSPNAYTLFS-ZINC22049444-1.pdb	-57.1273	-42.4123	-14.7149	0	20
VYKSPNAYTLFS-ZINC22049439-0.pdb	-53.5036	-41.2195	-12.6368	0.352678	18.9286
VYKSPNAYTLFS-ZINC22049416-0.pdb	-54.9318	-34.3585	-20.5733	0	21.1429
VYKSPNAYTLFS-ZINC22049412-0.pdb	-59.6842	-28.9907	-30.6936	0	19.5714
VYKSPNAYTLFS-ZINC22049403-0.pdb	-47.8125	-28.6341	-19.1784	0	25.875
VYKSPNAYTLFS-ZINC22049399-0.pdb	-53.5335	-28.1324	-25.4011	0	23.625
VYKSPNAYTLFS-ZINC22046554-0.pdb	-53.9482	-43.5389	-10.4093	0	18.4706
VYKSPNAYTLFS-ZINC22046550-1.pdb	-52.1847	-44.7578	-7.42684	0	19.3529
VYKSPNAYTLFS-ZINC22009511-0.pdb	-47.8225	-35.8296	-11.993	0	20.4545
VYKSPNAYTLFS-ZINC22002320-1.pdb	-58.8268	-46.0452	-12.7816	0	20.0667
VYKSPNAYTLFS-ZINC22002315-0.pdb	-50.4873	-38.2239	-12.2634	0	20
VYKSPNAYTLFS-ZINC22002311-1.pdb	-53.5317	-38.2134	-15.3184	0	18.9333
VYKSPNAYTLFS-ZINC22002306-0.pdb	-55.1233	-42.6433	-12.4799	0	19.8667
VYKSPNAYTLFS-ZINC21993116-1.pdb	-53.8751	-32.7426	-21.1325	0	21.8182
VYKSPNAYTLFS-ZINC21990362-1.pdb	-50.9287	-45.5609	-5.36783	0	19.0667

VYKSPNAYTLFS-ZINC21990359-1.pdb	-59.0912	-44.9025	-14.1887	0	19.5333
VYKSPNAYTLFS-ZINC21990356-0.pdb	-54.565	-52.065	-2.5	0	18.8235
VYKSPNAYTLFS-ZINC21990354-1.pdb	-58.4766	-52.4766	-6	0	18.4706
VYKSPNAYTLFS-ZINC21989403-1.pdb	-49.5877	-42.5877	-7	0	18.4667
VYKSPNAYTLFS-ZINC21989401-1.pdb	-49.7468	-42.7468	-7	0	18.4667
VYKSPNAYTLFS-ZINC21989398-1.pdb	-49.1666	-49.1666	0	0	18.6667
VYKSPNAYTLFS-ZINC21989396-1.pdb	-48.8228	-39.9819	-8.84095	0	18.5333
VYKSPNAYTLFS-ZINC21985602-0.pdb	-61.1329	-30.3022	-30.8307	0	19.0909
VYKSPNAYTLFS-ZINC21985597-1.pdb	-55.7793	-46.2793	-9.5	0	18.9286
VYKSPNAYTLFS-ZINC21985368-1.pdb	-46.8542	-45.7367	-1.11746	0	20.5
VYKSPNAYTLFS-ZINC21985366-1.pdb	-49.3807	-48.3366	-1.04408	0	21.2857
VYKSPNAYTLFS-ZINC20426942-1.pdb	-55.8461	-44.0007	-11.8455	0	19.7857
VYKSPNAYTLFS-ZINC20232286-0.pdb	-58.4186	-28.3646	-30.054	0	18.4615
VYKSPNAYTLFS-ZINC20232281-1.pdb	-64.0325	-34.3518	-29.6807	0	19.3077
VYKSPNAYTLFS-ZINC19907272-1.pdb	-51.2166	-44.2166	-7	0	19.3333
VYKSPNAYTLFS-ZINC19893242-0.pdb	-59.9221	-54.2077	-5.71446	0	19.2941
VYKSPNAYTLFS-ZINC19891524-1.pdb	-59.1218	-52.2758	-6.846	0	19.8824
VYKSPNAYTLFS-ZINC19872842-0.pdb	-48.0035	-39.5006	-8.5029	0	20.1667
VYKSPNAYTLFS-ZINC19872840-1.pdb	-49.6994	-43.5221	-6.17738	0	22.8333
VYKSPNAYTLFS-ZINC19872837-1.pdb	-57.0293	-51.9532	-5.07608	0	18.5
VYKSPNAYTLFS-ZINC17724845-0.pdb	-58.8854	-36.4528	-22.4326	0	18.7059
VYKSPNAYTLFS-ZINC17421447-1.pdb	-49.6549	-47.1549	-2.5	0	19
VYKSPNAYTLFS-ZINC17378346-0.pdb	-45.988	-28.8378	-17.1502	0	21.8
VYKSPNAYTLFS-ZINC17314801-1.pdb	-56.9317	-45.9317	-11	0	20.3571
VYKSPNAYTLFS-ZINC17300590-0.pdb	-57.9089	-41.2597	-16.6492	0	20.0769
VYKSPNAYTLFS-ZINC17288490-1.pdb	-54.8833	-45.3833	-9.5	0	21.7692
VYKSPNAYTLFS-ZINC17197391-0.pdb	-60.3834	-31.7644	-28.9382	0.319212	19
VYKSPNAYTLFS-ZINC17176129-0.pdb	-48.4032	-41.6407	-6.76253	0	19.5714
VYKSPNAYTLFS-ZINC17175346-0.pdb	-49.5346	-33.0346	-16.5	0	20
VYKSPNAYTLFS-ZINC17175261-1.pdb	-49.5511	-40.1477	-9.40337	0	24.1
VYKSPNAYTLFS-ZINC17173515-1.pdb	-54.6811	-34.4521	-20.5052	0.276203	23.8
VYKSPNAYTLFS-ZINC17164548-0.pdb	-57.6872	-51.9364	-5.75077	0	20.1667
VYKSPNAYTLFS-ZINC17161511-0.pdb	-49.6984	-37.0476	-12.6508	0	19.6923
VYKSPNAYTLFS-ZINC17108604-1.pdb	-57.3681	-39.5468	-17.8213	0	22.2727
VYKSPNAYTLFS-ZINC17064522-1.pdb	-57.9748	-50.9748	-7	0	19.6875
VYKSPNAYTLFS-ZINC17063817-0.pdb	-55.7344	-43.8036	-11.9309	0	20.0714

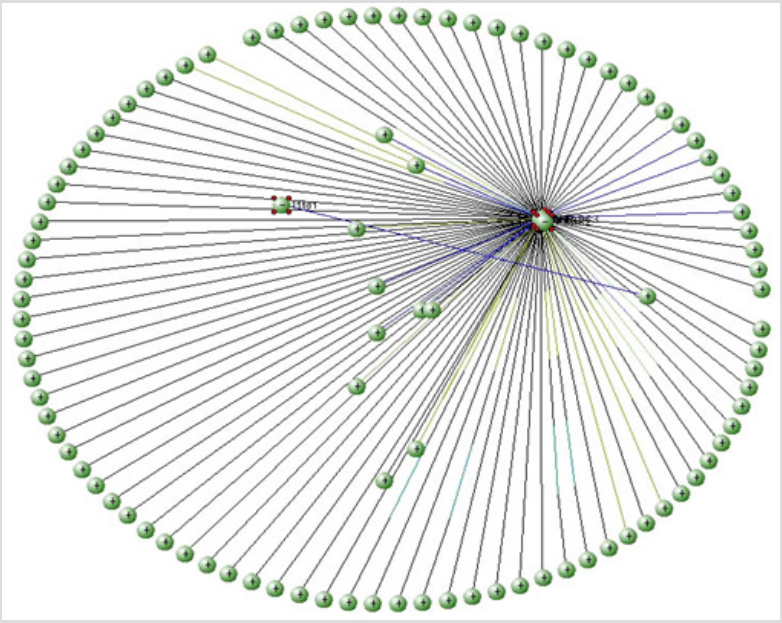


Figure 9: Hydrogen bonding contour (include H-bond acceptor (A) and H-bond donor (B) contour plots) of CoMSIA model.

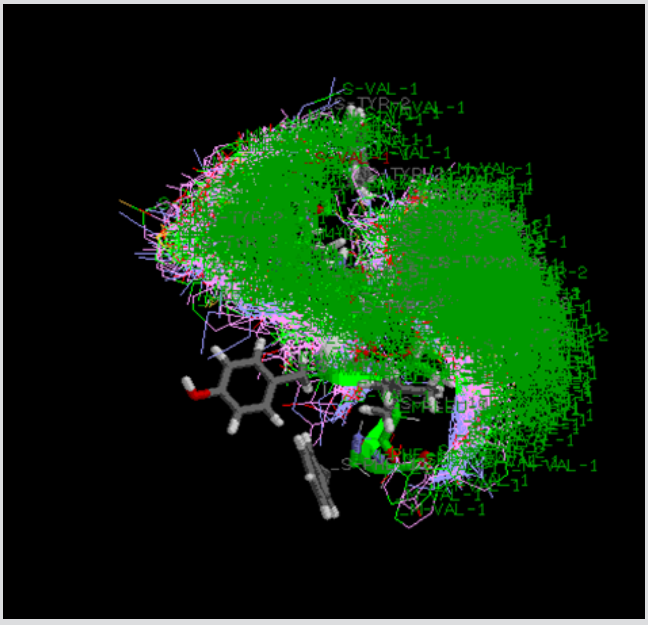


Figure 10

The results of the statistical analysis revealed that H-bond acceptor field had a significant contribution on the contour maps. The H-bond acceptor contour plots of the CoMSIA model containing the most active compound 28 were illustrated in Figure 9. The magenta contour around the N atom of the pyrazole ring indicated that the hydrogen bond acceptor group was conducive to the improvement of bioactivity. The nitrogen atom in this position may be beneficial to the generation of the hydrogen bond between the compound and its receptor protein [316]. It can also be proved by the fact that compounds 23,45,76, which displayed higher activities than other compounds. Thus, the H-bond acceptor group was believed to have a strong bioactivity on apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b. As shown in Figure 9, the regions in red implied where hydrogen bond donor group decrease activity. A large red tetrahedron surrounding the para-position of piperidine ring indicated that the hydrogen bond donor group was adverse to inhibitory activity, which can be demonstrated by the fact that compounds 11 and 13 showed a satisfactory activity. The other red contours were away from the most potent compound 28 and, hence, are not discussed. A prominent purple contour around the piperidine ring (R1 substitute) indicated that an H-bond donor was adverse to bioactivity Figure 10. For example, the order of many compounds' bioactivities was: 28 > 24 > 21. Compounds 23,45,76, with hydrogen bond donor groups showed low activities, which was in agreement with the contour map. The cyan color around the N-1 position of piperidine ring (R2 substitute) indicated that the hydrogen bond donor was favorable to activity. This result was the same to that of the hydrophobic contour plot. The magenta color shows the favored H-acceptor area, the red color shows the disfavored H-acceptor area, the cyan color shows the favored H-donor area, and the purple color represents the disfavored H-donor area (Figure 9,10).

According to the conclusions above, six new compounds (2a-2g) with satisfactorily predicted pIC50 values have been designed shown in Table 4. These designed compounds exhibited satisfactory predictive values also indicated the correctness of CoMFA and CoMSIA models (Table 4-6 and Figure 11-13).

Table 4:

VYKSPNAYTLFS-ZINC57678970-1.pdb	-58.5963	-19.6817	-38.9146	0	19
VYKSPNAYTLFS-ZINC57677644-0.pdb	-55.3012	-25.1254	-30.1757	0	21.2
VYKSPNAYTLFS-ZINC57677591-1.pdb	-57.3683	-45.3355	-12.0328	0	21.9286
VYKSPNAYTLFS-ZINC57477947-0.pdb	-59.8178	-49.3178	-10.5	0	18.4706
VYKSPNAYTLFS-ZINC57477944-1.pdb	-57.0127	-50.3768	-6.63595	0	18.5556
VYKSPNAYTLFS-ZINC57477942-1.pdb	-52.778	-50.278	-2.5	0	18.9333
VYKSPNAYTLFS-ZINC57312103-1.pdb	-53.9711	-43.4027	-10.5684	0	20.4286
VYKSPNAYTLFS-ZINC57218644-0.pdb	-58.9008	-41.4008	-17.5	0	19
VYKSPNAYTLFS-ZINC54186181-0.pdb	-46.7045	-36.5452	-10.1593	0	19
VYKSPNAYTLFS-ZINC50885362-1.pdb	-54.0129	-40.2304	-13.7825	0	21.5833
VYKSPNAYTLFS-ZINC49962147-0.pdb	-62.8932	-31.344	-31.5492	0	20
VYKSPNAYTLFS-ZINC49767070-0.pdb	-43.7908	-35.0951	-8.69569	0	20.1818
VYKSPNAYTLFS-ZINC49767069-1.pdb	-49.0032	-33.2887	-16.0264	0.31182	19.8182
VYKSPNAYTLFS-ZINC49757171-1.pdb	-51.6107	-32.1446	-19.466	0	19
VYKSPNAYTLFS-ZINC49757170-0.pdb	-46.3495	-35.6473	-10.7022	0	19.25
VYKSPNAYTLFS-ZINC49069304-0.pdb	-54.2836	-46.9519	-7.33167	0	18.5
VYKSPNAYTLFS-ZINC49069302-0.pdb	-58.0341	-37.0258	-21.0083	0	20.5625
VYKSPNAYTLFS-ZINC49069300-0.pdb	-51.6889	-41.8396	-9.84928	0	19.5
VYKSPNAYTLFS-ZINC49069298-0.pdb	-57.8139	-48.3139	-9.5	0	19.875
VYKSPNAYTLFS-ZINC48998271-1.pdb	-59.4902	-32.6852	-26.805	0	22.8333
VYKSPNAYTLFS-ZINC48998270-1.pdb	-50.9521	-38.9594	-11.9927	0	20
VYKSPNAYTLFS-ZINC48760380-1.pdb	-42.5037	-35.5037	-7	0	22.7778
VYKSPNAYTLFS-ZINC48760375-0.pdb	-58.399	-44.5706	-13.8284	0	22.0769
VYKSPNAYTLFS-ZINC47844082-0.pdb	-52.741	-46.4624	-6.27865	0	18.625
VYKSPNAYTLFS-ZINC45821218-1.pdb	-65.9396	-44.3974	-21.5423	0	24
VYKSPNAYTLFS-ZINC45352933-0.pdb	-52.4187	-22.7324	-29.6863	0	18.5385
VYKSPNAYTLFS-ZINC45336260-1.pdb	-46.3744	-39.3819	-6.99254	0	21.3
VYKSPNAYTLFS-ZINC45320233-1.pdb	-63.9813	-57.0222	-6.95908	0	20.1765
VYKSPNAYTLFS-ZINC45320229-0.pdb	-58.5018	-57.5036	-0.99818	0	19.8235
VYKSPNAYTLFS-ZINC45289072-1.pdb	-59.7213	-43.2674	-16.454	0	20.2
VYKSPNAYTLFS-ZINC45289069-1.pdb	-58.9176	-51.9176	-7	0	20.9333
VYKSPNAYTLFS-ZINC45286456-0.pdb	-58.9223	-43.2274	-15.6949	0	20.2667
VYKSPNAYTLFS-ZINC45286454-0.pdb	-53.5754	-44.205	-9.37042	0	18.6
VYKSPNAYTLFS-ZINC45285082-1.pdb	-54.8777	-52.7534	-2.12437	0	21.1333

Table 5:

VYKSPNAYTLFS-ZINC39132229-0.pdb	-46.3119	-24.7811	-21.5308	0	21.2222
VYKSPNAYTLFS-ZINC39132228-0.pdb	-45.7449	-35.2984	-10.4465	0	24.7778
VYKSPNAYTLFS-ZINC39128585-1.pdb	-55.4911	-42.5052	-12.9859	0	20.3077
VYKSPNAYTLFS-ZINC39128584-1.pdb	-55.0999	-40.6074	-14.4924	0	22.4615
VYKSPNAYTLFS-ZINC39127927-0.pdb	-56.3764	-36.4381	-20.273	0.33464	21.3333
VYKSPNAYTLFS-ZINC39126496-1.pdb	-69.9782	-40.079	-29.8993	0	19.8571
VYKSPNAYTLFS-ZINC39125943-0.pdb	-60.6704	-43.7195	-16.9509	0	20.3846
VYKSPNAYTLFS-ZINC39125942-1.pdb	-55.8554	-38.6813	-17.1741	0	19.2308
VYKSPNAYTLFS-ZINC39124748-1.pdb	-56.747	-54.2728	-2.47413	0	22.4
VYKSPNAYTLFS-ZINC39124686-0.pdb	-45.7651	-42.2651	-3.5	0	20.5
VYKSPNAYTLFS-ZINC39121706-1.pdb	-60.4509	-40.9178	-19.533	0	21.25
VYKSPNAYTLFS-ZINC39116592-0.pdb	-56.5597	-43.5097	-13.05	0	20.6923
VYKSPNAYTLFS-ZINC39074336-0.pdb	-60.7255	-39.0607	-21.6648	0	23
VYKSPNAYTLFS-ZINC38916803-0.pdb	-56.64	-47.1435	-9.49647	0	20.1429
VYKSPNAYTLFS-ZINC38887055-1.pdb	-44.9866	-30.9994	-13.9872	0	22.25
VYKSPNAYTLFS-ZINC38886715-0.pdb	-43.4995	-29.9403	-13.5592	0	22.375
VYKSPNAYTLFS-ZINC38881628-0.pdb	-47.3951	-40.3951	-7	0	22.4
VYKSPNAYTLFS-ZINC38864534-1.pdb	-49.5057	-36.118	-13.3877	0	20.9
VYKSPNAYTLFS-ZINC38812612-1.pdb	-60.8669	-51.3969	-9.47001	0	21
VYKSPNAYTLFS-ZINC38810455-0.pdb	-52.291	-41.8037	-10.4873	0	21.4615
VYKSPNAYTLFS-ZINC38810239-0.pdb	-55.8271	-34.1157	-21.7114	0	21.1667
VYKSPNAYTLFS-ZINC38599960-0.pdb	-65.0596	-48.5735	-16.7267	0.240602	19.4118
VYKSPNAYTLFS-ZINC38588766-0.pdb	-58.4844	-47.9844	-10.5	0	20.125
VYKSPNAYTLFS-ZINC38584800-0.pdb	-58.6568	-44.6568	-14	0	20.0714
VYKSPNAYTLFS-ZINC38529516-1.pdb	-48.6678	-34.6678	-14	0	22.1111
VYKSPNAYTLFS-ZINC38524652-1.pdb	-56.6553	-47.794	-8.86123	0	18.5
VYKSPNAYTLFS-ZINC38523994-0.pdb	-62.3042	-28.9477	-33.3564	0	19.5
VYKSPNAYTLFS-ZINC38350060-0.pdb	-45.9896	-36.5352	-9.45439	0	23.5556
VYKSPNAYTLFS-ZINC38283565-1.pdb	-61.2629	-44.8654	-16.3975	0	18.5333
VYKSPNAYTLFS-ZINC38265990-1.pdb	-57.3525	-43.3525	-14	0	20.8462
VYKSPNAYTLFS-ZINC38265497-0.pdb	-49.8455	-31.5683	-18.2773	0	20.3636
VYKSPNAYTLFS-ZINC38265336-1.pdb	-51.2343	-32.6022	-18.6321	0	21.4
VYKSPNAYTLFS-ZINC38265110-1.pdb	-37.9836	-30.9836	-7	0	20.5
VYKSPNAYTLFS-ZINC37637083-0.pdb	-51.0247	-32.9618	-18.3734	0.310532	18.4
VYKSPNAYTLFS-ZINC37632104-0.pdb	-46.0631	-39.0631	-7	0	19.6364
VYKSPNAYTLFS-ZINC37632102-0.pdb	-44.1426	-30.0537	-14.0889	0	19.1818

Table 6:

VYKSPNAYTLFS-ZINC37632098-0.pdb	-44.8618	-37.9151	-6.94668	0	22
VYKSPNAYTLFS-ZINC37632096-0.pdb	-52.4214	-52.4214	0	0	19.6
VYKSPNAYTLFS-ZINC37632094-1.pdb	-50.687	-47.187	-3.5	0	19.8667
VYKSPNAYTLFS-ZINC37625919-1.pdb	-51.8235	-43.7983	-8.02527	0	18.9286
VYKSPNAYTLFS-ZINC37624613-1.pdb	-57.0438	-44.5789	-12.4649	0	18.5294
VYKSPNAYTLFS-ZINC37624552-0.pdb	-57.9312	-41.6861	-16.2451	0	18.5
VYKSPNAYTLFS-ZINC37624550-1.pdb	-52.6843	-47.6912	-4.99309	0	20.6667
VYKSPNAYTLFS-ZINC37624549-0.pdb	-57.3588	-42.2554	-15.1034	0	21.2857
VYKSPNAYTLFS-ZINC37207007-0.pdb	-56.0689	-42.7788	-13.6337	0.343563	19.3571
VYKSPNAYTLFS-ZINC37207006-1.pdb	-50.0779	-42.1888	-7.88909	0	20.5714
VYKSPNAYTLFS-ZINC37207005-1.pdb	-46.7208	-37.9677	-8.75315	0	19.1538
VYKSPNAYTLFS-ZINC37207004-1.pdb	-54.0136	-48.7673	-5.24637	0	20.3333
VYKSPNAYTLFS-ZINC37207003-1.pdb	-49.621	-42.4247	-7.19635	0	20.4615
VYKSPNAYTLFS-ZINC37207002-1.pdb	-48.9212	-41.7939	-7.12733	0	21.3077
VYKSPNAYTLFS-ZINC37207001-1.pdb	-51.3067	-45.742	-5.56471	0	21.8462
VYKSPNAYTLFS-ZINC37207000-0.pdb	-49.0428	-45.6074	-3.43546	0	19.4615
VYKSPNAYTLFS-ZINC37206999-1.pdb	-50.5256	-48.0256	-2.5	0	21.3077
VYKSPNAYTLFS-ZINC37206998-0.pdb	-54.1823	-46.8566	-7.3257	0	20.5333
VYKSPNAYTLFS-ZINC37206997-1.pdb	-48.7175	-41.5643	-7.15321	0	21.3077
VYKSPNAYTLFS-ZINC37206996-0.pdb	-53.2583	-47.8997	-5.35856	0	21.4286
VYKSPNAYTLFS-ZINC37206995-0.pdb	-54.4435	-47.7214	-6.72207	0	20.1333
VYKSPNAYTLFS-ZINC37206994-1.pdb	-55.7661	-50.0448	-5.72133	0	20.1333
VYKSPNAYTLFS-ZINC37206993-1.pdb	-57.3039	-48.7359	-8.56804	0	21.4
VYKSPNAYTLFS-ZINC37206992-1.pdb	-54.8779	-45.7731	-9.10477	0	21.2143
VYKSPNAYTLFS-ZINC37206991-1.pdb	-50.7539	-48.2539	-2.5	0	21.8462
VYKSPNAYTLFS-ZINC37206990-0.pdb	-55.5462	-46.512	-9.03413	0	19.9333
VYKSPNAYTLFS-ZINC37206989-1.pdb	-51.2871	-48.7871	-2.5	0	21.8462
VYKSPNAYTLFS-ZINC37206988-0.pdb	-50.4755	-43.6367	-6.83878	0	20.3077
VYKSPNAYTLFS-ZINC37206987-1.pdb	-54.678	-40.7775	-13.9006	0	19.0714
VYKSPNAYTLFS-ZINC37206986-0.pdb	-51.3518	-45.0366	-6.31523	0	21.6923
VYKSPNAYTLFS-ZINC37206985-0.pdb	-52.6452	-44.2238	-8.42138	0	19.2667
VYKSPNAYTLFS-ZINC37206984-1.pdb	-51.8283	-43.038	-8.79036	0	21.8462
VYKSPNAYTLFS-ZINC37206983-0.pdb	-54.3013	-44.8013	-9.5	0	20.3571
VYKSPNAYTLFS-ZINC37206982-0.pdb	-52.4021	-48.5131	-3.88908	0	20.2857
VYKSPNAYTLFS-ZINC37206981-0.pdb	-51.0212	-44.7435	-6.27773	0	21.6923
VYKSPNAYTLFS-ZINC37206978-1.pdb	-46.57	-33.573	-12.9971	0	21.8182

VYKSPNAYTLFS-ZINC37206977-1.pdb	-51.6338	-45.573	-6.06082	0	21.9231
VYKSPNAYTLFS-ZINC37206976-0.pdb	-53.1776	-49.6776	-3.5	0	20
VYKSPNAYTLFS-ZINC37206975-0.pdb	-50.7849	-45.9576	-4.82735	0	18.6
VYKSPNAYTLFS-ZINC37206974-1.pdb	-55.4865	-45.9876	-9.49892	0	18.5
VYKSPNAYTLFS-ZINC37206973-1.pdb	-49.3602	-37.667	-11.6932	0	23.5455
VYKSPNAYTLFS-ZINC37206972-0.pdb	-56.6708	-40.5273	-16.1435	0	20.9333
VYKSPNAYTLFS-ZINC37206971-0.pdb	-47.9424	-35.3842	-12.8939	0.335697	19
VYKSPNAYTLFS-ZINC37206970-0.pdb	-49.672	-43.4366	-6.23544	0	22.6667
VYKSPNAYTLFS-ZINC37206969-0.pdb	-53.6235	-44.3033	-9.32019	0	21.7692
VYKSPNAYTLFS-ZINC37206968-0.pdb	-54.259	-48.0896	-6.16947	0	20.5333
VYKSPNAYTLFS-ZINC37206967-1.pdb	-49.0485	-43.0081	-6.04037	0	22.3333
VYKSPNAYTLFS-ZINC37206965-1.pdb	-50.34	-43.4585	-6.8815	0	20.3846
VYKSPNAYTLFS-ZINC37206964-1.pdb	-49.7352	-43.7858	-5.94947	0	22.8333
VYKSPNAYTLFS-ZINC37206963-0.pdb	-50.1802	-42.224	-7.95617	0	19.8462
VYKSPNAYTLFS-ZINC37206961-0.pdb	-51.423	-43.4679	-7.9551	0	20.7692
VYKSPNAYTLFS-ZINC37206960-0.pdb	-53.3327	-49.8327	-3.5	0	20.5
VYKSPNAYTLFS-ZINC37206959-1.pdb	-48.7473	-35.3685	-13.7397	0.360828	19
VYKSPNAYTLFS-ZINC37206958-0.pdb	-50.67	-42.2554	-8.41459	0	20.0769
VYKSPNAYTLFS-ZINC37206957-0.pdb	-54.7941	-46.2867	-8.50741	0	19.8
VYKSPNAYTLFS-ZINC37206956-0.pdb	-49.2597	-44.339	-4.92074	0	22.5833
VYKSPNAYTLFS-ZINC37206954-1.pdb	-52.0645	-45.8063	-6.2582	0	22.4615
VYKSPNAYTLFS-ZINC37206953-0.pdb	-60.7264	-47.0942	-13.6322	0	18.7059
VYKSPNAYTLFS-ZINC37206952-0.pdb	-49.2978	-41.0357	-8.26217	0	21.25
VYKSPNAYTLFS-ZINC37206951-0.pdb	-50.53	-43.8093	-6.72076	0	20.7692
VYKSPNAYTLFS-ZINC37206950-0.pdb	-44.7988	-30.3839	-14.778	0.363057	19.7
VYKSPNAYTLFS-ZINC37206949-0.pdb	-47.7996	-33.6115	-14.5434	0.355343	19.8182
VYKSPNAYTLFS-ZINC37206948-1.pdb	-50.1563	-43.8492	-6.30711	0	20.3846
VYKSPNAYTLFS-ZINC37206947-0.pdb	-52.2289	-44.0197	-8.20925	0	22.3846
VYKSPNAYTLFS-ZINC37206946-0.pdb	-51.5964	-43.8592	-7.73724	0	20.8462
VYKSPNAYTLFS-ZINC37206945-0.pdb	-51.1169	-33.5147	-17.6022	0	19.7273
VYKSPNAYTLFS-ZINC37206944-0.pdb	-44.6691	-29.9398	-15.0854	0.356065	20.1
VYKSPNAYTLFS-ZINC37206943-1.pdb	-54.343	-51.0946	-3.24836	0	20.6
VYKSPNAYTLFS-ZINC37206942-0.pdb	-49.5954	-47.4244	-2.17102	0	20.5385
VYKSPNAYTLFS-ZINC37206941-0.pdb	-54.6448	-37.2613	-17.6638	0.280307	18.7857
VYKSPNAYTLFS-ZINC37206940-1.pdb	-46.4157	-34.9883	-11.4274	0	22.1818
VYKSPNAYTLFS-ZINC37206939-1.pdb	-50.3485	-43.4961	-6.85247	0	20.2308

VYKSPNAYTLFS-ZINC37206938-0.pdb	-50.7434	-41.1268	-9.61659	0	21.8462
VYKSPNAYTLFS-ZINC37206937-1.pdb	-53.4754	-50.9754	-2.5	0	19.5333
VYKSPNAYTLFS-ZINC37206936-0.pdb	-50.8531	-48.3531	-2.5	0	20.2857
VYKSPNAYTLFS-ZINC37206934-1.pdb	-55.1908	-46.7479	-8.44292	0	19.4
VYKSPNAYTLFS-ZINC37206933-0.pdb	-58.3721	-49.8721	-8.5	0	19.0588
VYKSPNAYTLFS-ZINC37206932-1.pdb	-55.5742	-46.0742	-9.5	0	19.2
VYKSPNAYTLFS-ZINC37206931-1.pdb	-51.3911	-43.2765	-8.11456	0	20.7692
VYKSPNAYTLFS-ZINC37206930-0.pdb	-60.7277	-51.6594	-9.06827	0	20.7333
VYKSPNAYTLFS-ZINC37206929-0.pdb	-65.1279	-36.4358	-28.6921	0	19.6154
VYKSPNAYTLFS-ZINC37206928-0.pdb	-57.3459	-39.6533	-17.6926	0	23.7273
VYKSPNAYTLFS-ZINC37206927-0.pdb	-52.98	-36.71	-16.27	0	21.3333
VYKSPNAYTLFS-ZINC37206926-0.pdb	-57.8532	-40.927	-16.9262	0	22.6667
VYKSPNAYTLFS-ZINC37206925-0.pdb	-56.8439	-25.6243	-31.2197	0	19
VYKSPNAYTLFS-ZINC37206924-0.pdb	-59.1939	-36.0367	-23.1571	0	18.6429
VYKSPNAYTLFS-ZINC37206923-0.pdb	-52.5413	-42.2557	-10.2856	0	21.1667
VYKSPNAYTLFS-ZINC37206922-0.pdb	-58.6598	-46.336	-12.3238	0	21.4286
VYKSPNAYTLFS-ZINC37206921-0.pdb	-57.0063	-38.8567	-18.1496	0	23.7273
VYKSPNAYTLFS-ZINC37206920-0.pdb	-57.0996	-41.6796	-15.42	0	23.4167
VYKSPNAYTLFS-ZINC37206919-1.pdb	-63.3614	-47.0667	-16.2948	0	21.7143
VYKSPNAYTLFS-ZINC37206918-0.pdb	-68.2968	-36.2566	-32.0402	0	19.75
VYKSPNAYTLFS-ZINC37206917-0.pdb	-52.8841	-37.1631	-15.721	0	21.1667
VYKSPNAYTLFS-ZINC37206916-1.pdb	-55.8716	-32.1631	-23.7085	0	20.4615
VYKSPNAYTLFS-ZINC37206915-0.pdb	-57.2175	-25.6213	-31.5962	0	18.9091
VYKSPNAYTLFS-ZINC37206914-0.pdb	-63.2879	-47.9838	-15.3041	0	21.7143
VYKSPNAYTLFS-ZINC37206913-0.pdb	-71.126	-38.8954	-32.2306	0	18.7857
VYKSPNAYTLFS-ZINC37206912-0.pdb	-63.0284	-44.9629	-18.0655	0	20.2
VYKSPNAYTLFS-ZINC37206911-1.pdb	-68.4184	-38.0707	-30.3477	0	18.5333
VYKSPNAYTLFS-ZINC37206910-1.pdb	-61.7751	-30.3596	-31.4155	0	18.5385
VYKSPNAYTLFS-ZINC36889000-1.pdb	-53.2637	-39.4292	-13.8345	0	21.1667
VYKSPNAYTLFS-ZINC36867574-0.pdb	-62.147	-45.2553	-16.8917	0	22.9231
VYKSPNAYTLFS-ZINC36850489-1.pdb	-57.8183	-54.6767	-3.14161	0	19.5294
VYKSPNAYTLFS-ZINC36830580-1.pdb	-50.1224	-43.4017	-6.72067	0	20.2308
VYKSPNAYTLFS-ZINC36459927-1.pdb	-54.9887	-43.5223	-11.4665	0	21.6667
VYKSPNAYTLFS-ZINC36459147-1.pdb	-54.0961	-43.9694	-10.1267	0	22.5385
VYKSPNAYTLFS-ZINC36455624-0.pdb	-41.2127	-34.2127	-7	0	21.8889
VYKSPNAYTLFS-ZINC36420801-0.pdb	-51.0567	-37.928	-13.1287	0	21

VYKSPNAYTLFS-ZINC36420283-0.pdb	-65.4336	-48.0408	-17.3928	0	22.2857
VYKSPNAYTLFS-ZINC36420045-0.pdb	-49.8936	-41.7299	-8.16368	0	20.4615
VYKSPNAYTLFS-ZINC36381765-0.pdb	-53.8141	-49.2238	-4.59035	0	18.5
VYKSPNAYTLFS-ZINC36381764-1.pdb	-57.2869	-53.3854	-3.90145	0	18.8333
VYKSPNAYTLFS-ZINC36381341-0.pdb	-53.2417	-39.5838	-13.9074	0.24941	18.3333
VYKSPNAYTLFS-ZINC36379524-1.pdb	-65.2471	-32.9507	-32.2964	0	19.6364
VYKSPNAYTLFS-ZINC36379521-0.pdb	-55.7473	-40.7651	-14.9822	0	23.8182
VYKSPNAYTLFS-ZINC36379426-1.pdb	-46.5799	-39.5799	-7	0	20.2727
VYKSPNAYTLFS-ZINC36379425-1.pdb	-53.3486	-37.8486	-15.5	0	22.0909
VYKSPNAYTLFS-ZINC36379424-0.pdb	-51.7582	-40.0407	-11.7174	0	22.1818
VYKSPNAYTLFS-ZINC36379423-1.pdb	-47.934	-40.934	-7	0	21.1818
VYKSPNAYTLFS-ZINC36378525-0.pdb	-44.8157	-32.7979	-12.4344	0.416621	20.8
VYKSPNAYTLFS-ZINC36378524-1.pdb	-51.2297	-37.5103	-13.7195	0	24.6
VYKSPNAYTLFS-ZINC36378523-0.pdb	-51.0434	-30.7935	-20.2499	0	19.5
VYKSPNAYTLFS-ZINC36378522-0.pdb	-49.4292	-26.5149	-22.9142	0	21.4
VYKSPNAYTLFS-ZINC36377870-1.pdb	-57.5463	-39.7807	-17.7656	0	23.6364
VYKSPNAYTLFS-ZINC36330652-1.pdb	-54.9862	-44.9609	-10.0253	0	21.6
VYKSPNAYTLFS-ZINC36302664-1.pdb	-55.0862	-48.0862	-7	0	19.875
VYKSPNAYTLFS-ZINC36261825-0.pdb	-52.3552	-39.2962	-13.4019	0.342937	19
VYKSPNAYTLFS-ZINC35881877-1.pdb	-55.6377	-37.0725	-18.8289	0.263699	21.7273
VYKSPNAYTLFS-ZINC35854942-0.pdb	-51.7909	-34.8726	-16.9183	0	20.4167
VYKSPNAYTLFS-ZINC35854941-0.pdb	-56.9922	-38.2975	-18.6948	0	20.5833
VYKSPNAYTLFS-ZINC35822576-1.pdb	-47.5806	-34.4323	-13.1482	0	19
VYKSPNAYTLFS-ZINC35785168-1.pdb	-58.766	-35.9089	-23.1563	0.299232	18.4706
VYKSPNAYTLFS-ZINC35649603-1.pdb	-54.614	-41.113	-13.501	0	19.6429
VYKSPNAYTLFS-ZINC35636154-1.pdb	-62.2419	-43.7542	-18.4877	0	18.4706
VYKSPNAYTLFS-ZINC35636152-0.pdb	-57.8122	-43.6323	-14.1799	0	18.7647
VYKSPNAYTLFS-ZINC35606052-1.pdb	-48.4977	-41.4977	-7	0	19.8462
VYKSPNAYTLFS-ZINC35606049-0.pdb	-54.5027	-47.5027	-7	0	20.3846
VYKSPNAYTLFS-ZINC35606046-1.pdb	-45.5125	-39.8679	-5.64456	0	21.4615
VYKSPNAYTLFS-ZINC35606043-0.pdb	-47.34	-36.848	-10.492	0	19.0769
VYKSPNAYTLFS-ZINC35572481-0.pdb	-59.4402	-49.9402	-9.5	0	22.0714
VYKSPNAYTLFS-ZINC35572470-1.pdb	-54.3522	-47.3522	-7	0	19.6
VYKSPNAYTLFS-ZINC35566697-0.pdb	-52.9273	-45.9407	-6.98655	0	18.8
VYKSPNAYTLFS-ZINC35566695-0.pdb	-46.6688	-46.6688	0	0	18.9333
VYKSPNAYTLFS-ZINC35566693-1.pdb	-52.6698	-45.6698	-7	0	18.8667

VYKSPNAYTLFS-ZINC35566691-0.pdb	-49.292	-42.292	-7	0	18.6
VYKSPNAYTLFS-ZINC35394209-0.pdb	-53.2228	-41.253	-11.9698	0	20.7857
VYKSPNAYTLFS-ZINC35331398-1.pdb	-52.8795	-37.1585	-15.721	0	19.4615
VYKSPNAYTLFS-ZINC35308457-0.pdb	-54.1866	-47.3808	-6.80586	0	21.8571
VYKSPNAYTLFS-ZINC35302211-1.pdb	-58.535	-50.1726	-8.36243	0	18.5
VYKSPNAYTLFS-ZINC35285188-0.pdb	-57.1932	-41.0659	-16.1274	0	23.4167
VYKSPNAYTLFS-ZINC35271963-1.pdb	-59.4969	-56.0473	-3.44958	0	18.8333
VYKSPNAYTLFS-ZINC35260908-1.pdb	-60.9653	-47.1065	-13.8589	0	21.0667
VYKSPNAYTLFS-ZINC35260907-0.pdb	-54.6693	-47.4262	-7.24311	0	20.4667
VYKSPNAYTLFS-ZINC35260678-1.pdb	-60.4249	-44.5083	-15.9166	0	20.9286
VYKSPNAYTLFS-ZINC35246459-0.pdb	-51.4008	-42.9008	-8.5	0	18.4667
VYKSPNAYTLFS-ZINC35244485-1.pdb	-52.0903	-46.0903	-6	0	19.1429
VYKSPNAYTLFS-ZINC35244482-0.pdb	-50.1957	-43.1957	-7	0	20.8333
VYKSPNAYTLFS-ZINC35244481-0.pdb	-50.5068	-37.5104	-12.9964	0	18.6667
VYKSPNAYTLFS-ZINC35244480-1.pdb	-45.8798	-35.2663	-10.6136	0	20
VYKSPNAYTLFS-ZINC35244474-1.pdb	-54.3773	-48.5631	-5.81424	0	19.3333
VYKSPNAYTLFS-ZINC35244473-0.pdb	-48.5713	-45.0713	-3.5	0	20.2143
VYKSPNAYTLFS-ZINC35244472-1.pdb	-54.6464	-40.6464	-14	0	21.5455
VYKSPNAYTLFS-ZINC35244470-1.pdb	-47.3527	-42.9902	-4.3625	0	18.4615
VYKSPNAYTLFS-ZINC35244469-1.pdb	-54.1114	-50.6114	-3.5	0	18.5882
VYKSPNAYTLFS-ZINC35243471-1.pdb	-55.2692	-48.2692	-7	0	18.8
VYKSPNAYTLFS-ZINC35087834-1.pdb	-53.7837	-50.2837	-3.5	0	18.4706
VYKSPNAYTLFS-ZINC35087807-0.pdb	-56.0259	-46.5259	-9.5	0	20.2667
VYKSPNAYTLFS-ZINC35087805-0.pdb	-58.5557	-44.5557	-14	0	18.4667
VYKSPNAYTLFS-ZINC35077364-1.pdb	-57.3337	-45.7644	-11.5694	0	19.5
VYKSPNAYTLFS-ZINC35077363-1.pdb	-51.103	-41.7076	-9.39542	0	21.1429
VYKSPNAYTLFS-ZINC35077362-0.pdb	-52.8697	-46.3082	-6.56156	0	19.2941
VYKSPNAYTLFS-ZINC35077361-0.pdb	-49.7844	-42.7844	-7	0	18.4706
VYKSPNAYTLFS-ZINC35077360-1.pdb	-52.0429	-40.1874	-11.8555	0	23.2727
VYKSPNAYTLFS-ZINC35077359-1.pdb	-49.5434	-35.0119	-14.865	0.333444	19.8182
VYKSPNAYTLFS-ZINC35077358-0.pdb	-48.9921	-37.4414	-11.5508	0	18.4615
VYKSPNAYTLFS-ZINC35077357-0.pdb	-53.615	-37.865	-15.7501	0	21.9231
VYKSPNAYTLFS-ZINC35076132-1.pdb	-46.6477	-36.845	-9.80265	0	18.4706
VYKSPNAYTLFS-ZINC35076131-1.pdb	-53.323	-50.0434	-3.27956	0	19.5882
VYKSPNAYTLFS-ZINC35076130-0.pdb	-48.6348	-48.6348	0	0	18.4706
VYKSPNAYTLFS-ZINC35076129-0.pdb	-49.1304	-47.9184	-1.21204	0	18.9412

VYKSPNAYTLFS-ZINC35070559-0.pdb	-39.5979	-32.5979	-7	0	23.875
VYKSPNAYTLFS-ZINC35051407-1.pdb	-52.2454	-49.8954	-2.34997	0	20
VYKSPNAYTLFS-ZINC35051395-1.pdb	-48.3492	-48.3492	0	0	19.8667
VYKSPNAYTLFS-ZINC35051394-1.pdb	-58.7734	-52.9031	-5.87039	0	21.2667
VYKSPNAYTLFS-ZINC35051363-1.pdb	-53.2039	-42.7039	-10.5	0	20.3571
VYKSPNAYTLFS-ZINC35049526-1.pdb	-59.5657	-35.2666	-24.2991	0	18.5294
VYKSPNAYTLFS-ZINC35049524-1.pdb	-61.4284	-46.7195	-14.7089	0	20.1765
VYKSPNAYTLFS-ZINC35049105-1.pdb	-51.6859	-39.3339	-12.7577	0.405677	18.5714
VYKSPNAYTLFS-ZINC35035110-0.pdb	-55.942	-47.8078	-8.1342	0	19.8125
VYKSPNAYTLFS-ZINC35035101-0.pdb	-54.3607	-43.8664	-10.4943	0	20.6
VYKSPNAYTLFS-ZINC35035097-1.pdb	-54.0697	-50.5697	-3.5	0	20.2667
VYKSPNAYTLFS-ZINC35034960-0.pdb	-53.4309	-41.8604	-11.5705	0	19
VYKSPNAYTLFS-ZINC35033211-1.pdb	-57.029	-41.6074	-15.4217	0	21.4667

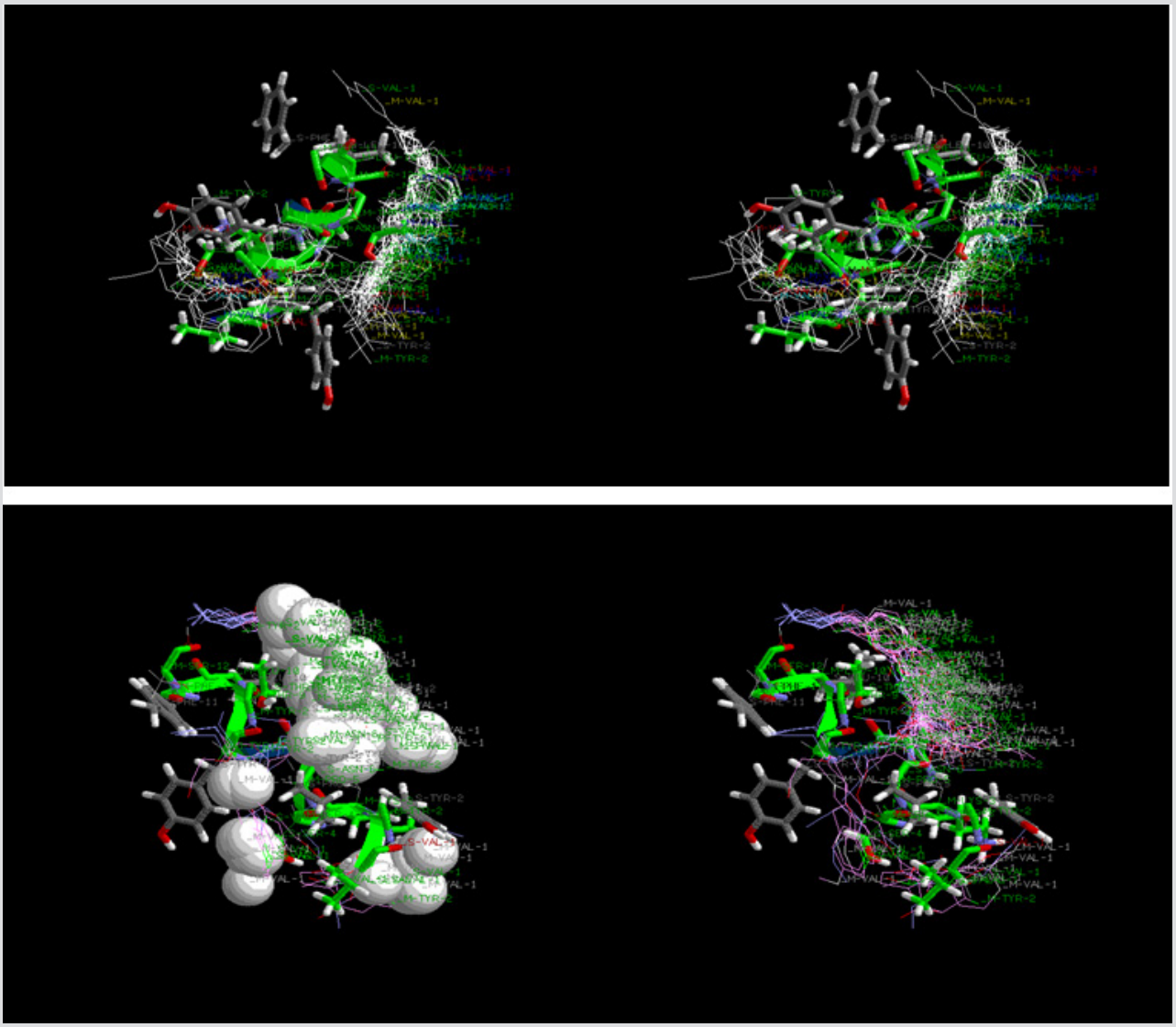


Figure 11

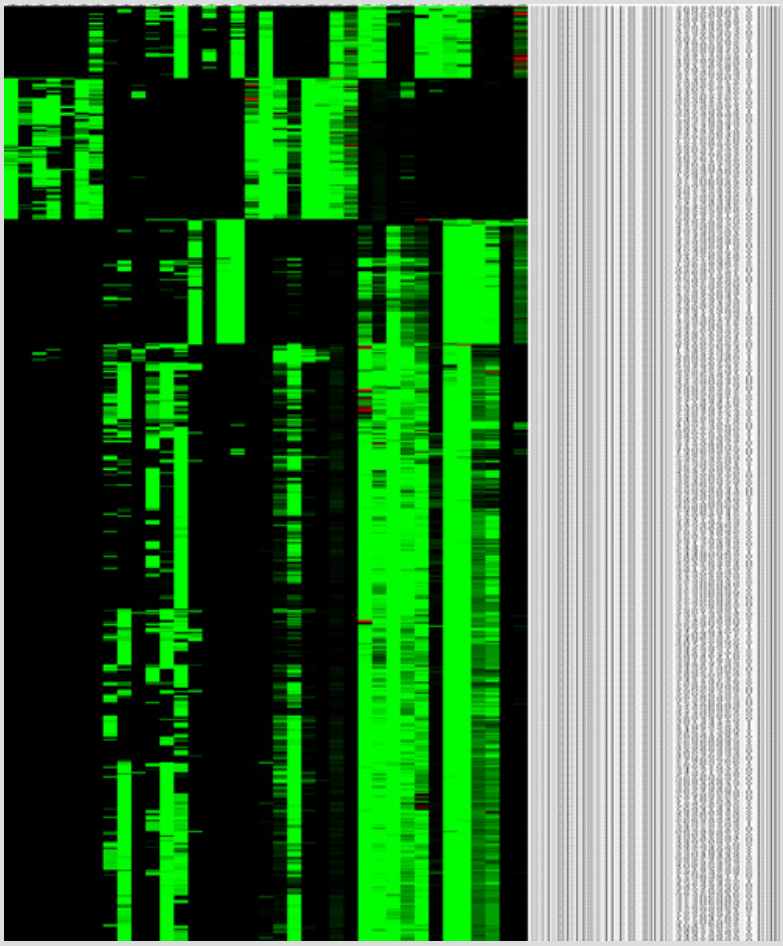


Figure 12: Designed molecules and predicted inhibit activities values of Chk 1 through CoMFA and CoMSIA. 23,45,76 model.

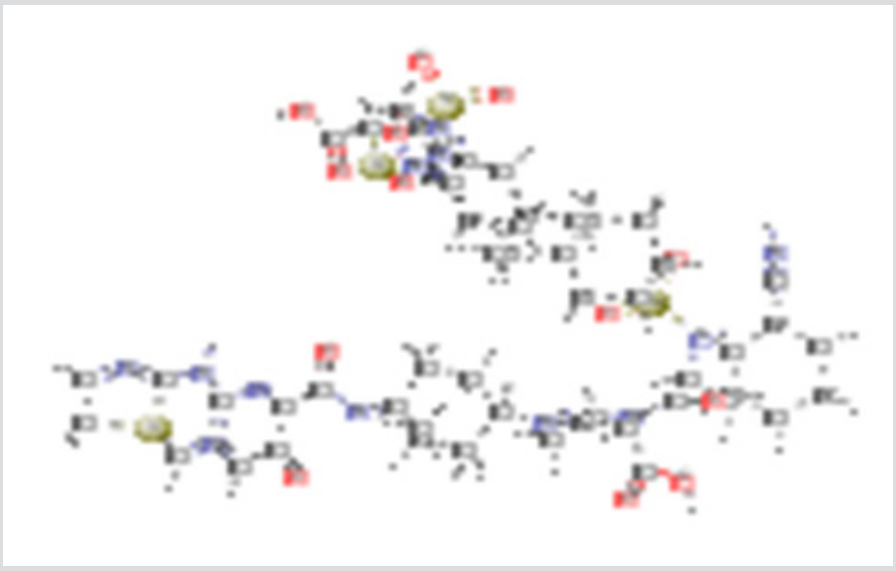


Figure 13

Molecular Docking

Molecular docking protocols were widely used to investigate the possible binding modes between the target derivatives and the receptor protein, which aided in the understanding the QSAR revealed by CoMFA/CoMSIA models. Prior to docking, a re-docking simulation was applied to validate the accuracy of molecular docking [317]. The target ligand taken from the crystal structure was re-docked into the active site with the root mean square deviation (RMSD) value of 0.225 Å for 4RVK. As shown in the Figure 14, the region of re-docking ligand (blue) was same to that of the original ligand (red). Three hydrogen bonds which were formed between the ligand and the protein appeared on the horizon. The H-bond distances were observed to be 1.69Å (C=O•••H-Cys87), 1.86Å(C=N•••H-Tyr86), 2.18Å (CN•••H-Lys38), respectively. In short, the surflex-dock program could successfully reproduce the original conformation. (Table 7) (Figure 14).

Table 7:

VYKSPNAYTLFS-ZINC34539694-1.pdb	-49.7451	-31.1403	-18.6049	0	22.1111
VYKSPNAYTLFS-ZINC34537793-1.pdb	-53.9113	-46.9113	-7	0	18.7333
VYKSPNAYTLFS-ZINC34537792-0.pdb	-51.6152	-42.1152	-9.5	0	18.6667
VYKSPNAYTLFS-ZINC34529374-0.pdb	-51.4012	-31.8079	-19.5933	0	22.4
VYKSPNAYTLFS-ZINC34529373-1.pdb	-50.4539	-24.5877	-25.8662	0	19
VYKSPNAYTLFS-ZINC34529002-1.pdb	-46.4895	-39.4895	-7	0	20.9091
VYKSPNAYTLFS-ZINC34529001-1.pdb	-43.3972	-37.6158	-5.78135	0	21.3636
VYKSPNAYTLFS-ZINC34527351-1.pdb	-64.4162	-41.7319	-22.6844	0	18.4444
VYKSPNAYTLFS-ZINC34516859-0.pdb	-60.1623	-48.0121	-12.1502	0	21.1429
VYKSPNAYTLFS-ZINC34514729-0.pdb	-45.7703	-38.7703	-7	0	22.9
VYKSPNAYTLFS-ZINC34514728-0.pdb	-45.9365	-38.9365	-7	0	24.2
VYKSPNAYTLFS-ZINC34514727-1.pdb	-44.509	-37.509	-7	0	23.7778
VYKSPNAYTLFS-ZINC34514726-0.pdb	-45.1049	-38.1094	-6.99556	0	23.3333
VYKSPNAYTLFS-ZINC34502526-0.pdb	-48.6625	-39.7033	-8.95924	0	18.5
VYKSPNAYTLFS-ZINC34500149-1.pdb	-63.6756	-51.129	-12.5467	0	19.3529

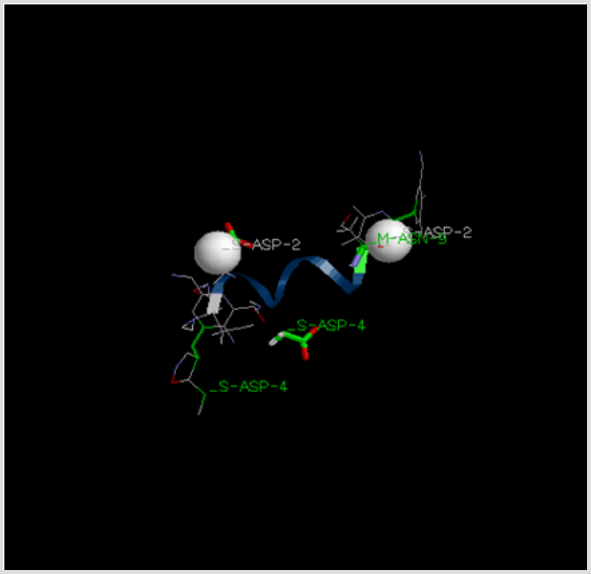


Figure 14: Re-docking results of the compound into the binding site of apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b protein and the crystal structure (blue ribbon) for the compound complex. The re-docked ligand (blue) and the original ligand (red) shown as a stick model.

The hydrogen-bonds are shown as yellow lines, with distance unit of Å.Subsequently, the most compound 28 was docked into the ligand-binding pocket of apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b

protein. As described in Figure 15, the docking results demonstrated three hydrogen bonds between compound 28 and the key residues (including Lys38, Glu85 and Cys87) in the Chk1 active pocket. The nitrogen atom of 1,7-diazacarbazole provided a hydrogen-bonding with Tyr86 (2.09Å), the N-H group was hydrogen bonded to residues Glu85 (2.06Å), and the 4-N of pyrazole ring formed a hydrogen-bonding with Lys38 (1.85Å). These residues could also interact with the target compound through electrostatic interactions. Moreover, the other residues, such as Tyr86, Glu91, Phe149, Asp148, and Glu55, could also stabilize ligand through electrostatic interaction [318]. The CoMFA/CoMSIA electrostatic-favorable blue contour around the pyrazole and the hetero atom further supported the structure-based analysis. In addition, Van der Waals forces were formed between the compound and important residues (for example: VYKSPNAYTLFS. This van der Walls interaction was involved in the activities of stabilizing compound in the active pocket. The docked model revealed that the hydrogen-bonding/electrostatic interactions played an important role in the interaction between the inhibitor and the protein, and the hydrogen bonds were similar to those in literature [17]. Moreover, the results obtained by the docking had been compared with the QSAR results to verify mutually. These interactions match well with the results of H-bond acceptor/electrostatic contour maps (Table 8) (Figure 15).

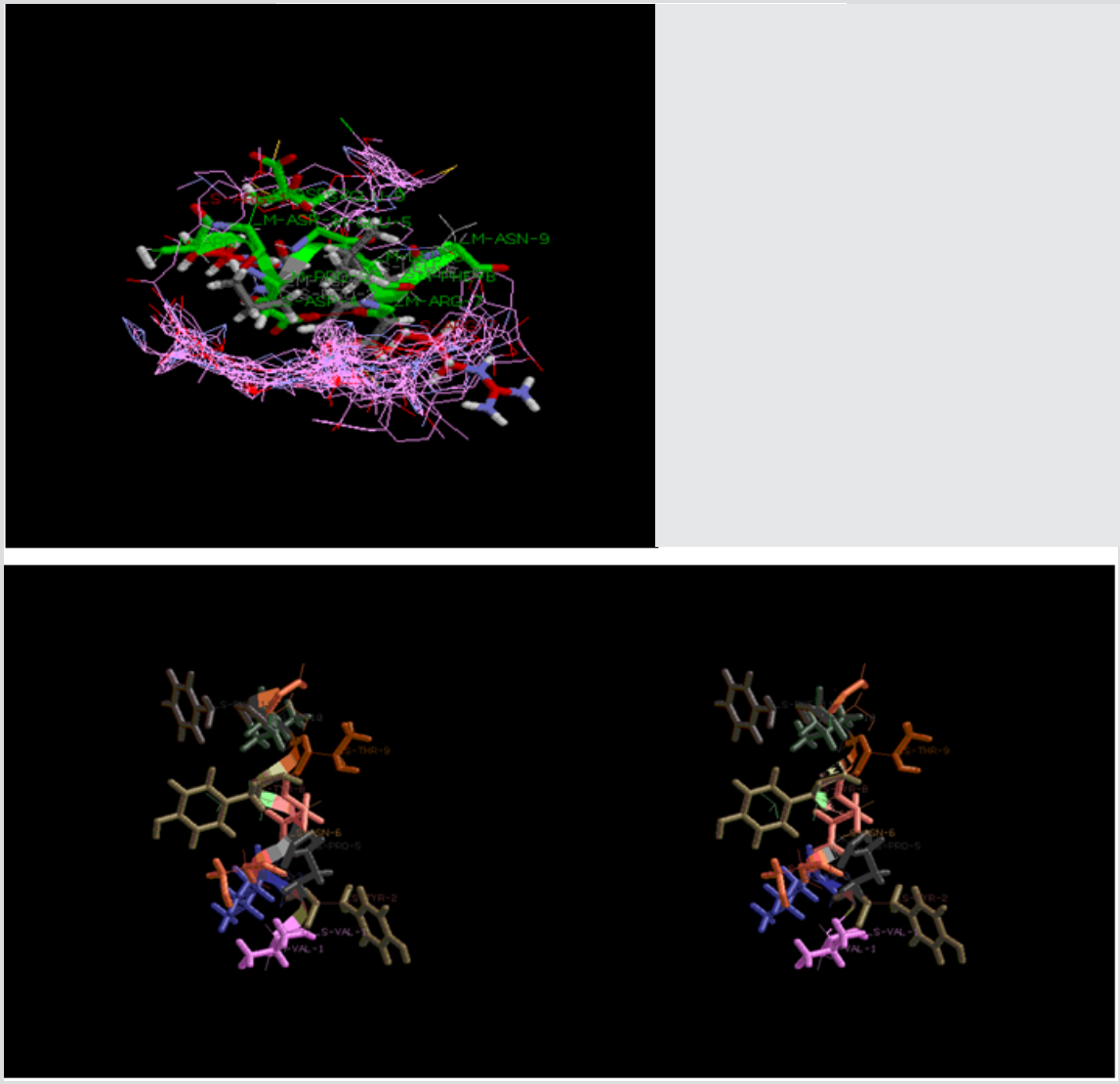


Figure 15: Docking result of compound 28 in the activity site of apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b protein. Ligand (yellow) and the key residues colored white for binding interaction were shown as stick model.

Table 8:

VYKSPNAYTLFS-ZINC34539694-1.pdb	-49.7451	-31.1403	-18.6049	0	22.1111
VYKSPNAYTLFS-ZINC34537793-1.pdb	-53.9113	-46.9113	-7	0	18.7333
VYKSPNAYTLFS-ZINC34537792-0.pdb	-51.6152	-42.1152	-9.5	0	18.6667
VYKSPNAYTLFS-ZINC34529374-0.pdb	-51.4012	-31.8079	-19.5933	0	22.4
VYKSPNAYTLFS-ZINC34529373-1.pdb	-50.4539	-24.5877	-25.8662	0	19

VYKSPNAYTLFS-ZINC34529002-1.pdb	-46.4895	-39.4895	-7	0	20.9091
VYKSPNAYTLFS-ZINC34529001-1.pdb	-43.3972	-37.6158	-5.78135	0	21.3636
VYKSPNAYTLFS-ZINC34527351-1.pdb	-64.4162	-41.7319	-22.6844	0	18.4444
VYKSPNAYTLFS-ZINC34516859-0.pdb	-60.1623	-48.0121	-12.1502	0	21.1429
VYKSPNAYTLFS-ZINC34514729-0.pdb	-45.7703	-38.7703	-7	0	22.9
VYKSPNAYTLFS-ZINC34514728-0.pdb	-45.9365	-38.9365	-7	0	24.2
VYKSPNAYTLFS-ZINC34514727-1.pdb	-44.509	-37.509	-7	0	23.7778
VYKSPNAYTLFS-ZINC34514726-0.pdb	-45.1049	-38.1094	-6.99556	0	23.3333
VYKSPNAYTLFS-ZINC34502526-0.pdb	-48.6625	-39.7033	-8.95924	0	18.5
VYKSPNAYTLFS-ZINC34500149-1.pdb	-63.6756	-51.129	-12.5467	0	19.3529

The hydrogen-bonds were described as yellow lines, with distance unit of Å.Compound 17 as a potential candidate was docked into the same active site to further analyze the impact of residues on the inhibitor activity. From the Figure 16, three hydrogen bonds were in sight. The H-bond distances were observed to be 1.88Å(C-N···H-Cys87), 2.03Å(N-H•••O-Glu85), and 2.18Å(CN•••H-N-Lys38), respectively. These H-bonds were the same to that of compound 28. Furthermore, electrostatic interaction was formed between this compound and the important residues (e.g., VYKSPNAYTLFS). These residues were mainly around R1substituent [319]. It could also confirm the importance of the electrostatic fields to bioactivity. Of course, the van der Waals force was essential. The amino residues, such as Ser147, Leu37, Val23, Glu91, Gly16, Asp94, Leu84, Val68, Ala36, Gly90, Ser88 and Gln13, could form van der Waals interaction with the target compound. These results also suggested that hydrogen bonding and electrostatic forces were the key interaction that confers bioactivity (Table 9 and Figure 16).

Table 9:

VYKSPNAYTLFS-ZINC34788430-1.pdb	-48.2371	-35.3442	-12.8928	0	22.4545
VYKSPNAYTLFS-ZINC34788427-0.pdb	-50.0478	-35.4176	-14.6302	0	18.6429
VYKSPNAYTLFS-ZINC34788423-0.pdb	-49.3081	-46.8081	-2.5	0	19.9286
VYKSPNAYTLFS-ZINC34787057-1.pdb	-51.94	-41.8296	-10.1105	0	19.4
VYKSPNAYTLFS-ZINC34776496-0.pdb	-68.9384	-36.97	-31.9683	0	18.6923
VYKSPNAYTLFS-ZINC34770532-0.pdb	-58.2325	-44.2325	-14	0	20.6154
VYKSPNAYTLFS-ZINC34770531-0.pdb	-58.3567	-44.3567	-14	0	20.8462
VYKSPNAYTLFS-ZINC34770257-0.pdb	-51.1638	-34.536	-16.6278	0	20.2143
VYKSPNAYTLFS-ZINC34768560-0.pdb	-45.8149	-37.3189	-8.49596	0	23.6
VYKSPNAYTLFS-ZINC34763656-0.pdb	-58.3077	-44.3077	-14	0	19.3571
VYKSPNAYTLFS-ZINC34763652-0.pdb	-51.1137	-44.1137	-7	0	20.9286
VYKSPNAYTLFS-ZINC34761710-1.pdb	-52.8246	-45.8246	-7	0	21.3571
VYKSPNAYTLFS-ZINC34759726-0.pdb	-48.2778	-41.2782	-6.99956	0	21.5833
VYKSPNAYTLFS-ZINC34759722-0.pdb	-51.2653	-40.2732	-10.9921	0	19.6667
VYKSPNAYTLFS-ZINC34753843-1.pdb	-65.3117	-46.4922	-18.8195	0	19.8125
VYKSPNAYTLFS-ZINC34753842-1.pdb	-58.4898	-40.4654	-18.0244	0	18.5
VYKSPNAYTLFS-ZINC34753733-0.pdb	-49.7142	-38.3415	-11.3728	0	20.2308

VYKSPNAYTLFS-ZINC34753615-1.pdb	-45.7788	-26.1664	-19.6123	0	26.4444
VYKSPNAYTLFS-ZINC34753614-1.pdb	-50.4741	-26.7953	-23.992	0.313162	24.3333
VYKSPNAYTLFS-ZINC34753360-0.pdb	-48.233	-44.733	-3.5	0	20.3333
VYKSPNAYTLFS-ZINC34744441-1.pdb	-52.8383	-36.8496	-15.9887	0	19.2941
VYKSPNAYTLFS-ZINC34744439-0.pdb	-64.5915	-41.3653	-23.2262	0	18.5294
VYKSPNAYTLFS-ZINC34744437-0.pdb	-58.8762	-35.8049	-23.0713	0	18.4706
VYKSPNAYTLFS-ZINC34744435-1.pdb	-57.0243	-41.9218	-15.1025	0	18.6471
VYKSPNAYTLFS-ZINC34729638-0.pdb	-52.497	-43.9952	-8.50181	0	19
VYKSPNAYTLFS-ZINC34729635-0.pdb	-53.4082	-46.283	-7.12513	0	21
VYKSPNAYTLFS-ZINC34628799-0.pdb	-58.8135	-48.742	-10.0715	0	21.6429
VYKSPNAYTLFS-ZINC34628320-0.pdb	-50.9941	-38.1996	-13.1623	0.367857	20.3077
VYKSPNAYTLFS-ZINC34628319-1.pdb	-48.7856	-39.2856	-9.5	0	18.4615
VYKSPNAYTLFS-ZINC34627996-0.pdb	-39.1611	-32.1611	-7	0	22
VYKSPNAYTLFS-ZINC34627152-0.pdb	-55.3854	-45.9079	-9.47743	0	18.7059
VYKSPNAYTLFS-ZINC34626586-0.pdb	-50.5478	-47.0478	-3.5	0	18.625
VYKSPNAYTLFS-ZINC34626582-1.pdb	-53.7004	-47.7004	-6	0	20.0667
VYKSPNAYTLFS-ZINC34624614-1.pdb	-51.2203	-44.9342	-6.28609	0	19.9231
VYKSPNAYTLFS-ZINC34621787-0.pdb	-49.6183	-41.7989	-7.81944	0	19
VYKSPNAYTLFS-ZINC34617645-1.pdb	-51.5264	-35.3286	-16.4466	0.248732	23.1
VYKSPNAYTLFS-ZINC34616245-1.pdb	-54.6167	-51.1167	-3.5	0	19.1875
VYKSPNAYTLFS-ZINC34614593-1.pdb	-43.5229	-29.5229	-14	0	22.7143
VYKSPNAYTLFS-ZINC34613363-0.pdb	-48.6226	-43.6228	-4.99979	0	22.0833
VYKSPNAYTLFS-ZINC34611341-0.pdb	-51.0215	-36.4185	-14.603	0	18.5714
VYKSPNAYTLFS-ZINC34611340-1.pdb	-50.5536	-47.0665	-3.48712	0	20.9286
VYKSPNAYTLFS-ZINC34611339-0.pdb	-58.3178	-45.1636	-13.1542	0	23
VYKSPNAYTLFS-ZINC34611338-0.pdb	-57.0789	-40.4489	-16.63	0	18.5
VYKSPNAYTLFS-ZINC34611007-0.pdb	-54.846	-46.1104	-8.73566	0	19.875
VYKSPNAYTLFS-ZINC34601304-1.pdb	-54.4359	-40.4359	-14	0	20.7273
VYKSPNAYTLFS-ZINC34601303-1.pdb	-50.2116	-36.2239	-13.9877	0	21.2727
VYKSPNAYTLFS-ZINC34596136-0.pdb	-44.0042	-27.0042	-17	0	22.8571
VYKSPNAYTLFS-ZINC34592649-1.pdb	-43.766	-33.266	-10.5	0	20.3
VYKSPNAYTLFS-ZINC34590642-0.pdb	-57.2038	-39.3999	-17.8039	0	23.7273
VYKSPNAYTLFS-ZINC34579789-1.pdb	-62.9102	-46.1194	-16.7908	0	20.4118

VYKSPNAYTLFS-ZINC34579209-0.pdb	-45.0762	-25.7236	-19.3526	0	23.75
VYKSPNAYTLFS-ZINC34579208-1.pdb	-45.4169	-29.9432	-15.4736	0	24
VYKSPNAYTLFS-ZINC34579143-0.pdb	-48.6282	-28.5429	-20.0853	0	26
VYKSPNAYTLFS-ZINC34579142-1.pdb	-45.7017	-25.351	-20.3506	0	23.5
VYKSPNAYTLFS-ZINC34574889-1.pdb	-50.6556	-38.1462	-12.5093	0	18.5
VYKSPNAYTLFS-ZINC34574888-0.pdb	-60.9554	-46.9554	-14	0	18.5625
VYKSPNAYTLFS-ZINC34574680-1.pdb	-60.1131	-50.6453	-9.46779	0	20.1875
VYKSPNAYTLFS-ZINC34571646-0.pdb	-54.8258	-40.8309	-13.9949	0	19.3571
VYKSPNAYTLFS-ZINC34569717-0.pdb	-47.717	-38.217	-9.5	0	19.6364
VYKSPNAYTLFS-ZINC34567952-1.pdb	-56.8868	-47.1544	-9.73243	0	19.625
VYKSPNAYTLFS-ZINC34567947-1.pdb	-54.6815	-39.8485	-14.833	0	18.5294
VYKSPNAYTLFS-ZINC34567945-0.pdb	-51.0024	-42.4719	-8.5305	0	18.8125
VYKSPNAYTLFS-ZINC34567935-1.pdb	-59.8674	-40.0729	-19.7945	0	18.75
VYKSPNAYTLFS-ZINC34567932-0.pdb	-58.4604	-46.881	-11.5794	0	19.1765
VYKSPNAYTLFS-ZINC34567930-1.pdb	-47.7422	-30.0813	-18.0965	0.435573	18.4375
VYKSPNAYTLFS-ZINC34567927-1.pdb	-46.9936	-41.9991	-4.9945	0	18.625
VYKSPNAYTLFS-ZINC34567673-0.pdb	-52.0067	-43.6633	-8.34344	0	21.3077
VYKSPNAYTLFS-ZINC34567670-1.pdb	-53.5203	-43.9656	-9.55463	0	22.3846
VYKSPNAYTLFS-ZINC34567373-1.pdb	-58.0468	-34.1427	-23.9041	0	21.5833
VYKSPNAYTLFS-ZINC34567104-1.pdb	-53.4235	-32.5896	-20.834	0	20.5
VYKSPNAYTLFS-ZINC34561654-1.pdb	-51.7971	-33.0231	-18.774	0	19.8182
VYKSPNAYTLFS-ZINC34559924-0.pdb	-46.0665	-32.7299	-13.3367	0	21.2727
VYKSPNAYTLFS-ZINC34557241-1.pdb	-42.3828	-35.45	-6.93278	0	19
VYKSPNAYTLFS-ZINC34557240-0.pdb	-43.6055	-33.1055	-10.5	0	20.0833
VYKSPNAYTLFS-ZINC34557239-1.pdb	-40.8927	-33.9598	-6.93289	0	19
VYKSPNAYTLFS-ZINC34557238-0.pdb	-42.4653	-35.4653	-7	0	19.3636
VYKSPNAYTLFS-ZINC34552885-1.pdb	-45.3933	-38.4076	-6.98569	0	20.7273
VYKSPNAYTLFS-ZINC34552848-0.pdb	-57.3892	-44.5961	-12.7931	0	20.0667
VYKSPNAYTLFS-ZINC34552406-1.pdb	-54.7481	-45.2481	-9.5	0	21
VYKSPNAYTLFS-ZINC34551303-0.pdb	-54.48	-41.1148	-13.3652	0	20.1538
VYKSPNAYTLFS-ZINC34541600-0.pdb	-52.8983	-39.0032	-13.8951	0	18.6429
VYKSPNAYTLFS-ZINC34541048-1.pdb	-58.0057	-48.6838	-9.32193	0	19
VYKSPNAYTLFS-ZINC34541046-1.pdb	-60.3863	-46.3863	-14	0	18.5294

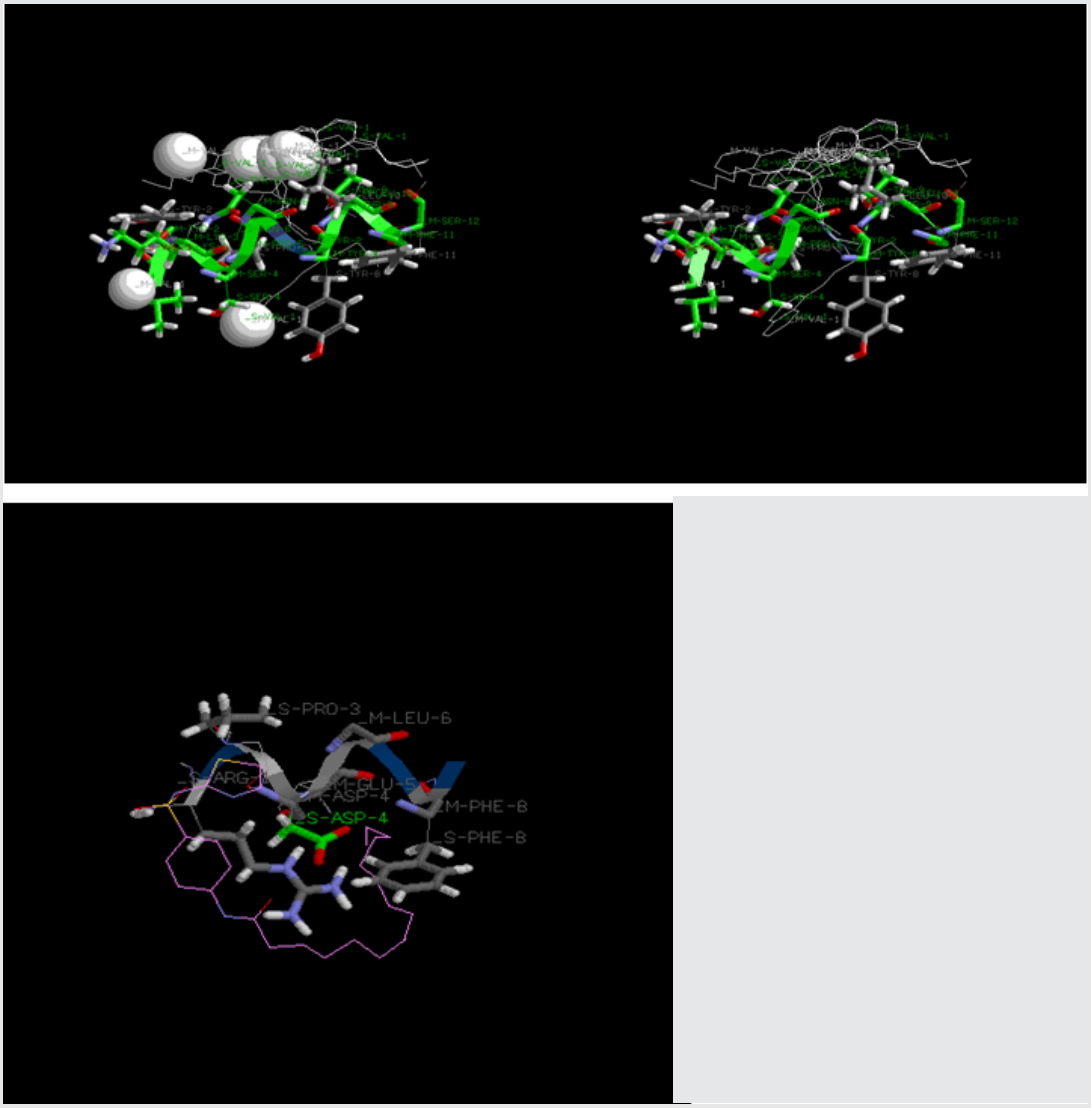


Figure 16: Docking result of compound 17 in the activity site of apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b protein. Ligands (yellow) and the key residues (white) for binding interactions were shown as a stick model.

The hydrogen bonds were described as yellow lines, with distance measured in Å.

Molecular Dynamics Simulations (Figure 17)

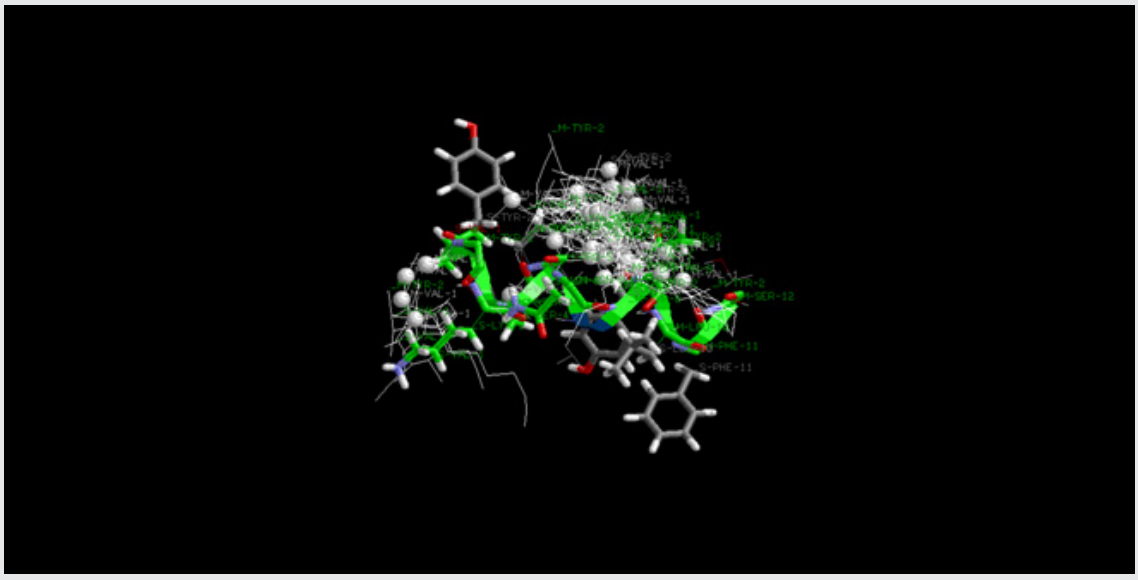


Figure 17

In a subsequent step, the molecular dynamics process was performed on compounds 28 and 17 to further explore the probable binding modes between the compounds and the receptor protein. Complex Chk1 structure of the compound as generated by the previous docking modelling was used as the initial coordinates for molecular dynamics study. The molecular dynamics simulations were performed using the dynamics module of SYBYL-X 2.0 at the vacuum environment [40,41]. Then, the energy minimization was performed for the complex molecule with Gasteiger-Hückel charge and Tripos force field without water using Boltzmann initial velocity. The simulations were executed using normal temperature and volume (NTV) [42] ensemble 300 K with coupling 100 fs. Additionally, we perform a 5 ns simulation with a time step of 1 fs and snapshot the conformation every 1000 fs.

A 5 ns simulation of the complex between the compound 28 and the protein was run to energy balance at 2 ns to obtain the stable conformation. The total energy of compound ranging from 4057 to 3620 KJ/mol was illustrated in Figure 18. After 2 ns, the total energy of the complex dropped to 3620 KJ/mol, and tended to stability. This result suggested that the ligand-protein complex could reach the metastable conformation after 2 ns of simulation [320-322]. The alignment of original and molecular dynamics simulated ligand was shown in Figure 19 to indicate a high similarity among these two ligand. From the Figure 19, some residues (such as: Tyr86, Ser147, Gly90, Asp94, Asp148, Glu17, Leu15, Leu137, Leu84, and Gly16) were still important to the interaction between compound 28 and Chk1 protein. However, the number of hydrogen bonds and the amino acid residues decreased. One hydrogen bond was formed between Gly89 and compound 28, with the distance of 1.67Å. This reside formed not only hydrogen-bonding to ligand, but also formed the electrostatic force. Moreover, most residues (such as: Asp148, Glu17, Leu15, et al.) were bound to the target compound through the electrostatic interaction. Therefore, we predict that the electrostatic interaction has a greater impact on activity of compound 28. These results match well with the field contribution of CoMSIA model (Figure 18,19).

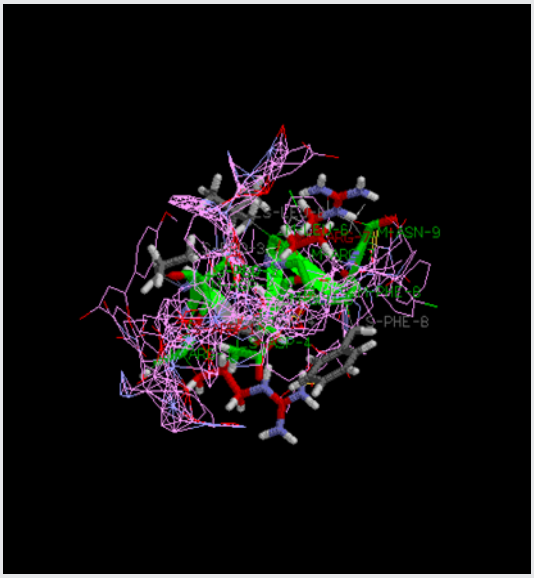


Figure 18: Plot of the total-energy docked complex (compound 28 (A) and compound 17 (B)) and the molecular dynamics simulation time (ps) in the molecular dynamics simulated structure.

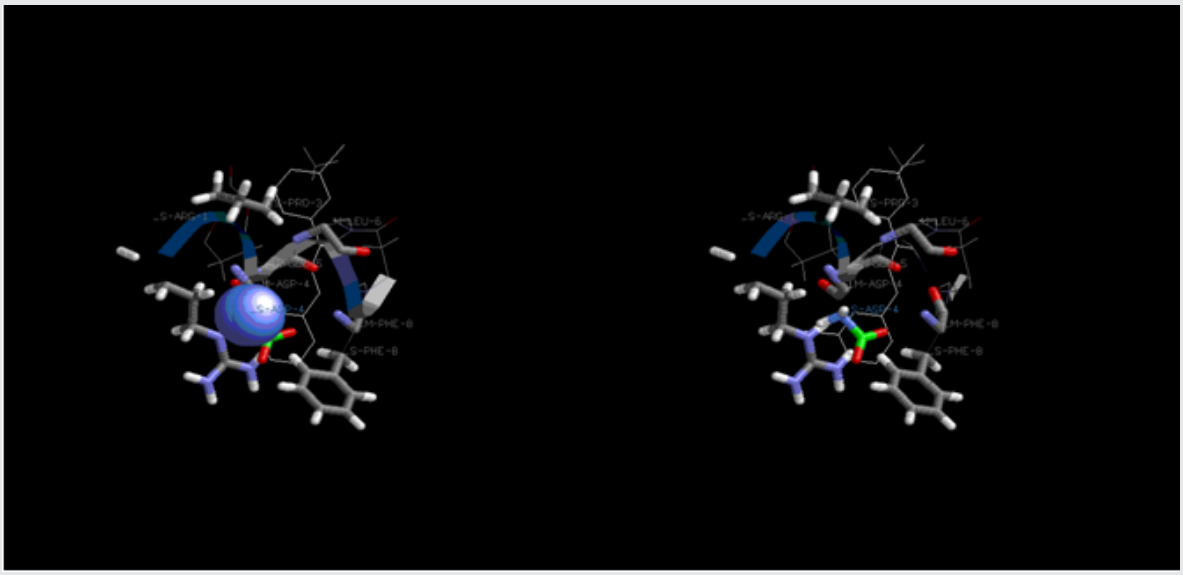


Figure 19: The alignment of molecular dynamics simulated and original ligand, as well as the docking results of compound 28 (A) and 17 (B).

The docked complex of compound 17 and the Chk1 protein was studied subsequently through a 5 ns simulation. The total energy of the docked complex was shown in Figure 18. From Figure 18, the total energy tends to be stable after 1ns. The alignment of molecular dynamics simulated and the original ligand was described in Figure 19. Meanwhile, the key residues which interacted with the target compound were in sight. Some key residues (Glu85, Cys87, and Lys38) still interacted with compounds through H-bonding and electrostatic force. The H-bond distances were 1.87Å, 1.95Å, and 1.86Å, respectively. Combining the docking results, the key residues (Glu85, Cys87, and Lys38) were predicted to be the important factor influencing the bioactivity of apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b.

Numerical results

In this section, two benchmark test problems, whose exact solutions are known, are solved by the proposed combination of cubic spline and explicit group iterative method [323]. The results are then compared with those obtained by the combination of cubic spline with block Gauss-Seidel iterative method (SBGS)

Combination of central difference scheme with explicit group iterative method (CDEG) where the CDEG scheme can be derived by substituting the partial derivative in Eq. (1) by the central difference approximation. In all cases, we assume that u(0) = 0 as the initial guess and the iterations were stopped when the estimated error was below tolerance, that is when $|u(s+1)-u(s)| \leq 10\text{-}12$ was achieved.

Example 1

Consider the convection-diffusion equation

$\partial 2u\partial x2+\partial 2u\partial y2=\beta\partial u\partial x,0<x,y<1$

The exact solution for the problem is given by

Computational Complexity Analysis

In order to show the efficiency of the proposed method, computational complexity of the SEG iterative method is examined. Assume that the solution domain is discretized into even intervals, Nx and Ny in x- and y-directions, respectively. Therefore, we have (nx-1) (ny-1) grouped points and (nx+ny-1) ungrouped points, where nx = Nx-1 and ny = Ny-1. This can be shown as in Figure 20.

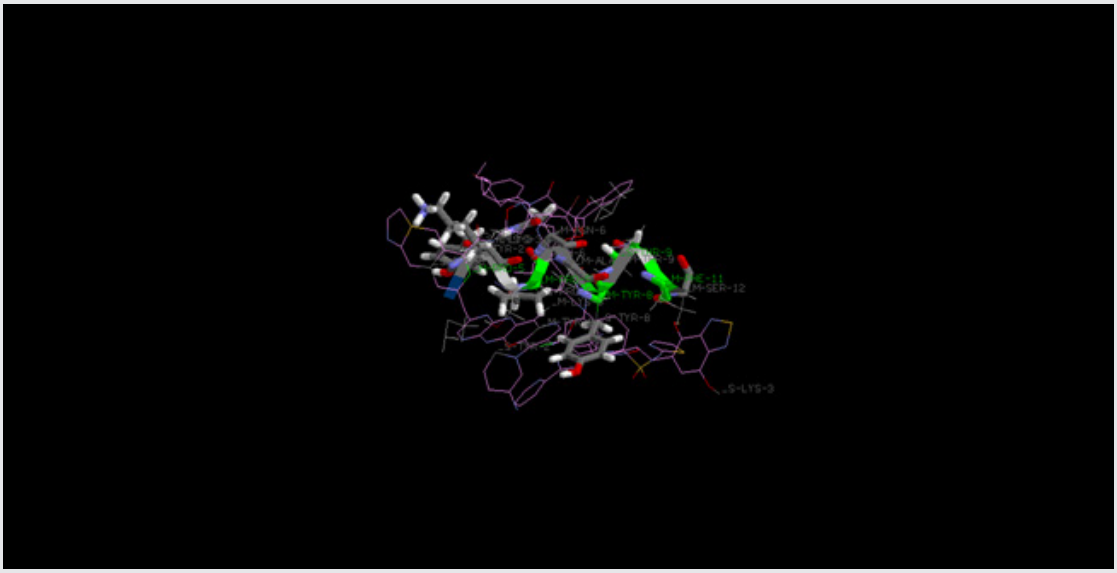


Figure 20

Types of points in SEG for Nx = Ny = 8.

The estimation on this computational complexity is based on the arithmetic operations performed at each titeration for the additions/substractions (Add/Sub) and multiplications/divisions (Mul/Div) operations. Therefore, the number of operations required for SEG is given as in Table 1. The total number of arithmetic operations can be obtained by multiplying the number of arithmetic operations for each titeration with the number of titerations.

The number of arithmetic operations per iteration for SEG iterative method

Numerical Results: In this section, two benchmark test problems, whose exact solutions are known, are solved by the proposed combination of cubic spline and explicit group iterative method. The results are then compared with those obtained by the combination of cubic spline with block Gauss-Seidel iterative method (SBGS) [324]. Combination of central difference scheme with explicit group iterative method (CDEG) where the CDEG scheme can be derived by substituting the partial derivative in Eq (1) by the central difference approximation. In all cases, we assume that u(0) = 0 as the initial guess and the iterations were stopped when the estimated error was below tolerance, that is when $|u(s+1) - u(s)| \leq 10\text{-}12$ was achieved.

Example 1

Consider the convection-diffusion equation

∂2u∂x2+∂2u∂y2=β∂u∂x,0<x,y<1

The exact solution for the problem is given by

u(x,y)=eβx2sinπysinhσ[2e-β2sinhσx+sinhσ(1-x)]

where σ2=π2+β24 and β > 0. The boundary conditions can be obtained from the exact solution. The (4 × 4) matrix system can be obtained by substituting D00 = β, D10 = D20 = 0 and Gl, m = Gl+1, m = Gl+1, m+1 = Gl, m+1 = 0 in Eq (20). .

Computational errors for proposed method, SEG compared with CDEG [24] and SBGS [29] by using β = 10 and k/h2 = 64.

Total arithmetic operations needed to generate the above results for CDEG [24], SBGS [29] and the proposed method, SEG (β = 10, k/h2 = 64).

Example 2

Given the following Poisson's equation in polar cylindrical coordinates in r - z plane.

∂2u∂r2+∂2u∂z2+1r∂u∂r=coshz (5rcoshr+2(2+r2)sinhr),0<r,z<1

The exact solution is u(r, z) = r2 sinh r cosh z. The solutions can be approximated by replacing the variables (x, y) by (r, z) and substituting D(r) =-1r and g(r, z) = cosh z(5r cosh r + 2(2 + r2) sinh r) into scheme (20).

Maximum absolute errors of proposed method, SEG compared with CDEG [24] and SBGS [29] compared to the exact solution (k/h = 0.8).

Total arithmetic operations needed to generate the above results for CDEG [24], SBGS [29] and the proposed method, SEG for k/h = 0.8.

Algorithm

In this section, we will design a sub-algorithm and the main algorithm, respectively. These two algorithms are listed as follows.

Initial point algorithm (sub-algorithm)

Step 0: Given any x0 ∈ Rn, δ1, δ2 ∈ (0, 1), ϵk > 0, r ∈ (0, 1), ϵ ∈ (0, 1), let k: = 0.

Step 1: If ‖ ek ‖ ≤ ϵ, stop. Otherwise let dk = -ek and go to next step.

Step 2: Choose ϵk+1 satisfies (1.4) and let αk = 1, r, r2, r3, … until (1.3) holds.

Step 3: Let xk+1 = xk + αkdk.

Step 4: If ‖ ek+1 ‖ ≤ ϵ, stop.

Step 5: Compute dk+1 = -ek+1 + βkdk, set k = k + 1 and go to Step 2.

Remark

(i) βk of Step 5 is a scalar and different βk will determine different CG methods.

(ii) From Step 2 and [34], it is easy to deduce that there exists αk such that (1.3). Thus, this sub-algorithm is well defined.

In the following, we will state the main algorithm. First, assume that the terminated point of sub-algorithm is xsup, and then the given algorithm is defined as follows.

Algorithm 1

Main algorithm

Step 0: Choose xsup ∈ Rn as the initial point, an initial symmetric positive definite matrix B0 ∈ Rn×n, and constants r, σ ∈ (0, 1), ϵmain < ϵ, a positive integer M > 0, m(k) = 0, let k: = 0;

Step 1: Stop if ‖ ek ‖ ≤ ϵmain. Otherwise solve (1.10) to get dk.

Step 2: Let αk = 1, r, r2, r3,… until (1.13) holds.

Step 3: Let the next iterative be xk+1 = xk + αkdk.

Step 4: Update Bk by quasi-Newton updates formula and ensures the update matrix Bk+1 is positive definite.

Step 5: Let k: = k + 1. Go to Step 1.

Remark

Step 4 of Algorithm 1 can ensure that Bk is always positive definite. This means that (1.10) has a unique solution dk. By positive definiteness of Bk, it is easy to obtain ekTdk<0. In the following sections, we only concentrate to the convergence of the main algorithm.

Convergence analysis

Let Ω be the level set with

$\Omega = \{x| \| e(x) \| \leq \| e(x0) \| \}$.-03.1

Similar to [37,38,56], the following assumptions are needed to prove the global convergence of Algorithm 1. Assumption A

(i) e is continuously differentiable on an open convex set Ω1 containing Ω.

(ii) The Jacobian of e is symmetric, bounded, and positive definite on Ω1, i.e., there exist positive constants $M* \geq m* > 0$ such that

$\| \nabla e(x) \| \leq M* \forall x \in \Omega 1$-3.2

and

$m*\|d\|2 \leq dT\nabla e(x)d \forall x \in \Omega 1, d \in Rn$.-3.3

Assumption B

Bk is a good approximation to ∇ek, i.e.,

$\| (\nabla ek-Bk)dk \| \leq \epsilon* \|ek\|$,-3.4

where $\epsilon* \in (0, 1)$ is a small quantity.

Considering Assumption B and using the von Neumann lemma, we deduce that Bk is also bounded (see [37]).

Lemma 3.1

Let Assumption B hold. Then dk is a descent direction of p(x) at xk, i.e.,

$\nabla p(xk)Tdk \leq -(1-\epsilon*)\|e(xk)\|2$.-3.5

Proof

By using (1.10), we get

$\nabla p(xk)Tdk=e(xk)T\nabla e(xk)dk=e(xk)T[(\nabla e(xk)-Bk)dk-e(xk)]=e(xk)T(\nabla e(xk)-Bk)dk-e(xk)Te(xk)$.-3.6

Thus, we have

$\nabla p(xk)Tdk+\|e(xk)\|2=e(xk)T(\nabla e(xk)-Bk)dk \leq \|e(xk)\|\|(\nabla e(xk)-Bk)dk\|$.

It follows from (3.4) that

$\nabla p(xk)Tdk \leq \|e(xk)\|\|(\nabla e(xk)-Bk)dk\|-\|e(xk)\|2 \leq -(1-\epsilon*)\|e(xk)\|2$.-3.7

The proof is complete.

The following lemma shows that the line search technique (1.13) is reasonable, and then Algorithm 1 is well defined.

Lemma 3.2

Let Assumptions A&B hold. Then Algorithm 1 will produce an iteration $xk+1 = xk + \alpha kdk$ in a finite number of backtracking steps.

Proof

From Lemma 3.5 in [38] we have in a finite number of backtracking steps

$p(xk + \alpha kdk) \leq p(xk) + \alpha k\sigma e(xk)Tdk$,

from which, in view of the definition of $p(xl(k)) = max0 \leq j \leq m(k)\{p(xk-j)\} \geq p(xk)$, we obtain (1.13). Thus we conclude the result of this lemma. The proof is complete.

Now we establish the global convergence theorem of Algorithm 1.

Theorem 3.1

Let Assumptions A and B hold, and {αk, dk, xk+1, ek+1} be generated by Algorithm 1. Then

$\lim k \to \infty \|ek\|=0$.-3.8

Proof

By the acceptance rule (1.13), we have

p(xk+1)-p(xl(k))≤σαkekTdk<0.-3.9

Using m(k + 1) ≤ m(k) + 1 and p(xk+1) ≤ p(xl(k)), we obtain

p(xl(k+1)) ≤ max {p(xl(k)), p(xk+1)} = p(xl(k)).

This means that the sequence {p(xl(k))} is decreasing for all k. Then {p(xl(k))} is convergent. Based on Assumptions A and B, similar to Lemma 3.4 in [32], it is not difficult to deduce that there exist constants b1 ≥ b2 > 0 such that

b2‖dk‖2≤dkTBkdk=-ekTdk≤b1‖dk‖2.-3.10

By (1.13) and (3.10), for all k > M, we get

p(xl(k))=p(xl(k)-1+αl(k)-1dl(k)-1)≤max0≤j≤m(l(k)-1){p(xl(k)-j-1)}+σαl(k)-1gl(k)-1Tdl(k)-1≤max0≤j≤m(l(k)-1){p(xl(k)-j-1)}-σb2αl(k)-1‖dl(k)-1‖2.-3.11

Since {p(xl(k))} is convergent, from the above inequality, we have

limk→∞αl(k)-1‖dl(k)-1‖2=0.

This implies that either

limk→∞infdl(k)-1=0-3.12

or

limk→∞infαl(k)-1=0.-3.13

If (3.12) holds, following [40], by induction we can prove that

limk→∞‖dl(k)-j‖=0-3.14

and

limk→∞p(xl(k)-j)=limk→∞p(xl(k))

for any positive integer j. As k ≥ l(k) ≥ k - M and M is a positive constant, by

xk = xk-M-1 + αk-M-1dk-M-1 + ⋯ + αl(k)-1dl(k)-1

and (3.14), it can be derived that

limk→∞p(xl(k))=limk→∞p(xk).-3.15

According to (3.10) and the rule for accepting the step αkdk,

p(xk+1)-p(xl(k))≤αkσekTdk≤αkσb2‖dk‖2.-3.16

This means

limk→∞αk‖dk‖2=0,

which implies that?

limk→∞αk=0-3.17

or

limk→∞‖dk‖=0.-3.18

If equation (3.18) holds, since Bk is bounded, then ‖ ek ‖ = ‖ Bkdk ‖ ≤ ‖ Bk ‖ ‖ dk ‖ → 0 holds. The conclusion of this lemma holds. If (3.17) holds. Then acceptance rule (1.13) means that, for all large enough k, αk′=αkr such that

p(xk+αk′dk)-p(xk)≥p(xk+αk′dk)-p(xl(k))>σαk′ekTdk.-3.19

Since

p(xk+αk′dk)-p(xk)=αk′∇p(xk)Tdk+o(αk′‖dk‖).-3.20

Using (3.19) and (3.20) in [38], we have

∇p(xk)Tdk=ekT∇e(xk)dk≤δ∗ekTdk,

where δ∗ > 0 is a constant and σ < δ∗. So we get

[δ∗-σ]αk′ekTdk+o(αk′‖dk‖)≥0.-3.21

Note that δ∗ - σ > 0 and ekTdk<0, we have from dividing (3.21) by αk′‖dk‖

limk→∞ekTdk‖dk‖=0.-3.22

By (3.10), we have

limk→∞‖dk‖=0.-3.23

Consider ‖ ek ‖ = ‖ Bkdk ‖ ≤ ‖ Bk ‖ ‖ dk ‖ and the bounded Bk again, we complete the proof.

Lemma 3.3

see Lemma 4.1 in [37]

Let e be continuously differentiable, and ∇e(x) be nonsingular at x∗ which satisfies e(x∗) = 0. Let

a≡{‖∇e(x∗)‖+12c,2c},c=‖∇e(x∗)-1‖.-3.24

If ‖ xk - x∗ ‖ sufficiently small, then the inequality

1a‖xk-x∗‖≤‖e(xk)‖≤a‖xk-x∗‖-3.25

holds.

Theorem 3.2

Let the assumptions in Lemma 3.3 hold. Assume that there exists a sufficiently small ε0 > 0 such that ‖ Bk - ∇e(xk) ‖ ≤ ε0 for each k. Then the sequence {xk} converges to x∗ superlinearly for αk = 1. Moreover, if e is q-order smooth at x∗ and there is a neighborhood U of x∗ satisfying for any xk ∈ U,

‖ [Bk - ∇e(x∗)](xk - x∗) ‖ ≤ η‖xk-x∗‖1+q,-3.26

Since g is continuously differentiable and ∇e(x) is nonsingular at x∗, there exists a constant γ > 0 and a neighborhood U of x∗ satisfying

max { ‖ ∇e(y) ‖ , ‖ ∇e(y)-1 ‖ } ≤ γ,

where ∇e(y) is nonsingular for any y ∈ U. Consider the following equality when αk = 1:

Bk(xk+1-x∗)+[∇e(xk)(xk-x∗)-Bk(xk-x∗)]+[e(xk)-e(x∗)-∇e(xk)(xk-x∗)]=e(xk)+Bkdk=0,-3.27

the second term and the third term are o(‖ xk - x∗ ‖). By the von Neumann lemma, and considering that ∇e(xk) is nonsingular, Bk is also nonsingular. For any y ∈ U and ∇e(y) being nonsingular and max { ‖ ∇e(y) ‖ , ‖ ∇e(y)-1 ‖ } ≤ γ, then we obtain from Lemma 3.3

‖ xk+1 - x∗ ‖ = o(‖ xk - x∗ ‖) = o(‖ e(xk) ‖), as k → ∞,

this means that the sequence {xk} converges to x∗ superlinearly for αk = 1.

If e is q-order smooth at x∗, then we get

e(xk)-e(x∗) - ∇e(xk)(xk - x∗) = O(‖xk-x∗‖q+1).

Consider the second term of (3.27) as xk → x∗, and use (3.26), we can deduce that the second term of (3.27) is also O(‖xk-x∗‖q+1). Therefore, we have

‖ xk+1 - x∗ ‖ = O(‖xk-x∗‖q+1), as xk → x∗.

The proof is complete. □Redundant short-time Fourier methods, also known as Gabor analysis [1], are widely used in signal processing applications. The basic idea is the analysis of a signal f by consideration of the projections ⟨f,gτ,ω⟩ of f onto time-frequency atoms gτ,ω. The gτ,ω are obtained by translation of a unique prototype function over time and frequency: gτ,ω(t)=g(t-τ)e2πitω. This classical construction leads to a signal decomposition with fixed time-frequency resolution over the whole time-frequency plane. The restriction to a fixed resolution is often undesirable in processing signals with variable time-frequency characteristics. Alternative decompositions have been introduced to overcome this deficit, e.g. the wavelet transform [2], the constant-Q transform (CQT) [3] or decompositions using filter banks [4], in particular based on perceptive frequency scales [5]. Adaptation over time is considered in approaches such as modulated lapped transforms [6], adapted local trigonometric transforms [7] or (time-varying) wavelet packets [8].

Most of the cited work achieves flexible tilings of the time-frequency plane, but efficient reconstruction from signal-adaptive, overcomplete time-frequency transforms is rarely addressed. One exception is a recent approach in [9], which is in fact a special case of the more general model considered in the present paper. The wealth of existing approaches to fast adaptive transforms underlines the need for flexibility arising from many applications. On the other hand, the introduction of flexibility in a transform that is based on accurate mathematical modeling causes technical complications that are not always easy to overcome. We introduce an approach to fast adaptive time-frequency transforms that is based on a generalization of painless nonorthogonal expansions [10]. It allows for adaptivity of the analysis windows and the sampling points. Since the resulting frames locally resemble classical Gabor frames and share some of their structure, they are called nonstationary Gabor frames.

Efficiency: The computation time of the nonstationary Gabor transform was found to be better than a recent fast CQT implementation [30], as seen in Figure 20. The two plots show mean values for computation time in seconds and the corresponding variance over 50 iterations, with varying window lengths and number of frequency bins, respectively. The outlier, drawn in gray, in Figure 20 (left) at the prime number 600 569 illustrates dependence of the current CQ-NSGT implementation on the signal length's prime factor structure, analogous to FFT (Figure 20).

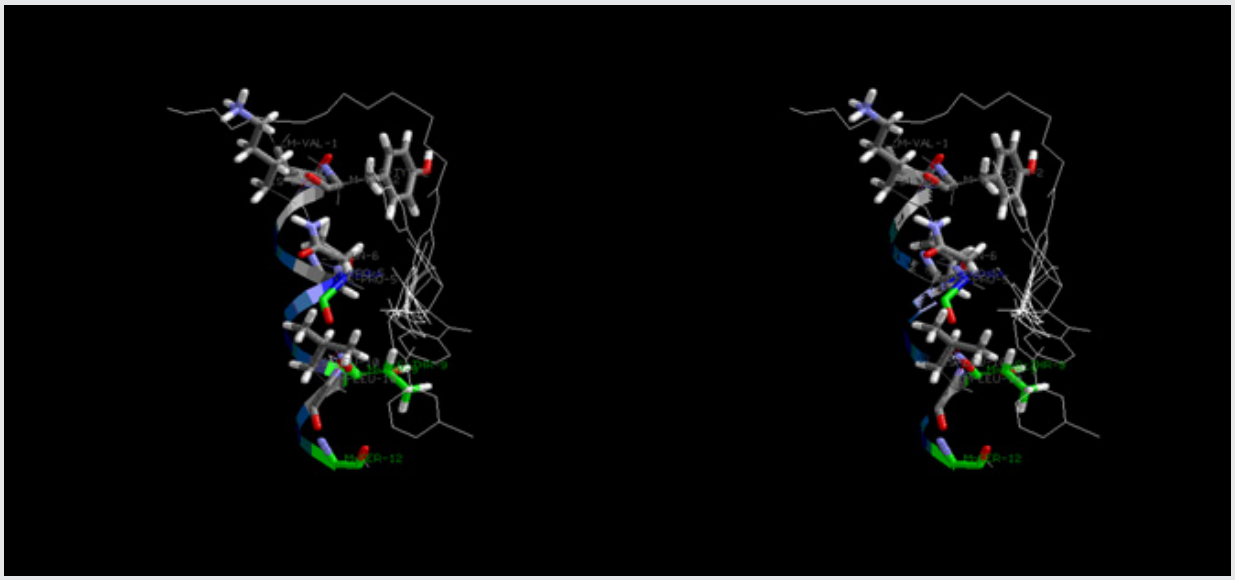


Figure 21

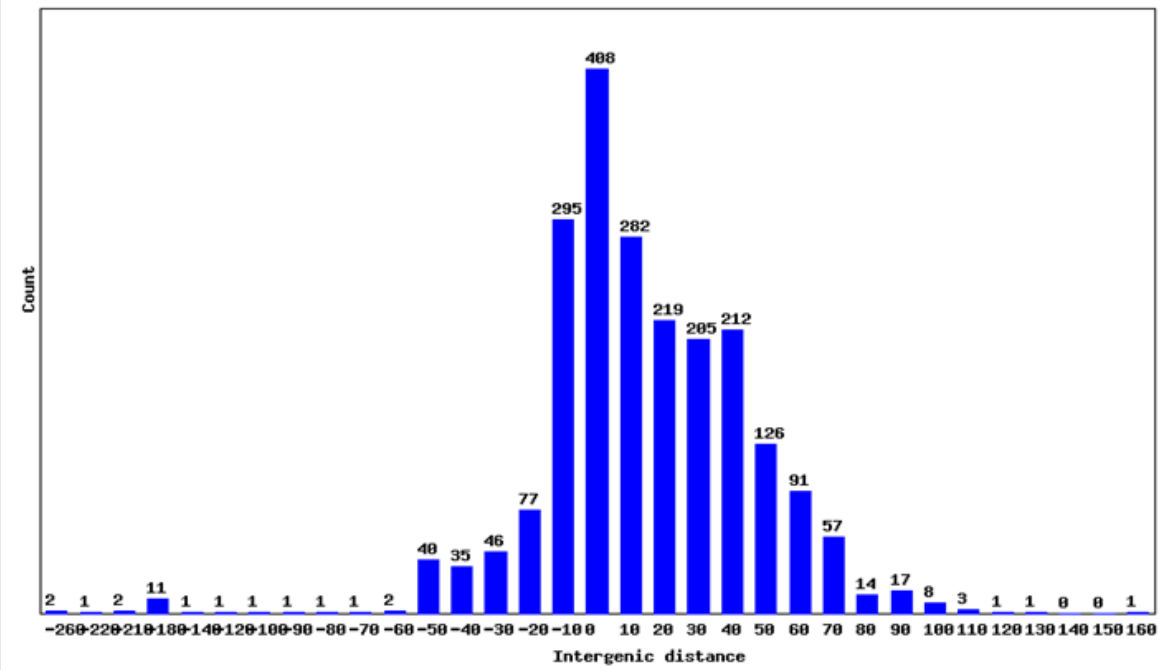


Figure 22

Glockenspiel (Example 1) Regular Gabor, constant-Q nonstationary Gabor and constant-Q representations of the in silico Reaction-diffusion Chemoproteomic-aided Molecular designed, ovarian stem cells (OSCs) motif VYKSPNAYTLFS derived and a quasi-Newton algorithm Scan Predicted antigenic peptide-like polypharmacophoric CAMSPCPL) signal [325-328]. The transform parameters were B=48 and ξmin=200Hz (Figure 21). Regular and constant-Q nonstationary Gabor representations of the in silico Reaction-diffusion Chemoproteomic-aided Molecular designed, ovarian stem cells (OSCs) motif VYKSPNAYTLFS derived and a quasi-Newton algorithm Scan Predicted antigenic peptide-like polypharmacophoric signal. The transform parameters were B=48 and ξmin=75Hz (Figure 22). Regular and constant-Q nonstationary Gabor representations of the in silico Reaction-diffusion Chemoproteomic-aided Molecular designed, ovarian stem cells (OSCs) motif VYKSPNAYTLFS derived and a quasi-Newton algorithm Scan Predicted antigenic peptide-like polypharmacophoric signal. The transform parameters were B=48 and ξmin=50Hz (Figure 23).

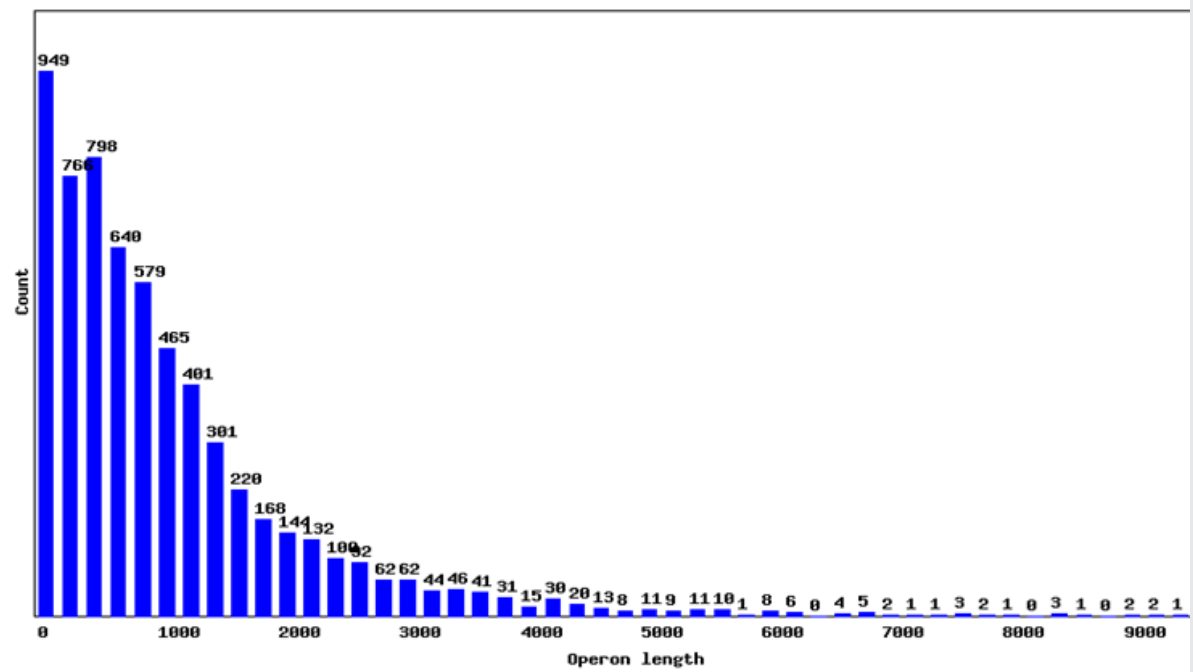


Figure 23

Comparison of computation time of CQT (top curves) and NSGT (bottom curves) on the in silico Reaction-diffusion Chemoproteomic-aided Molecular designed, ovarian stem cells (OSCs) motif VYKSPNAYTLFS derived and a quasi-Newton algorithm Scan Predicted antigenic peptide-like polypharmacophoric ligand (CAMSPCPL) targeted on conserved (β2GPI) antibodies, Anti-phospholipid antibodies (aPL) or from (PAPP-A; gi:38045915) and APOH apolipoprotein H post-trancripts domains . The Figure 23 on the left shows the computation times for signals of various lengths with the number of bins per octave fixed at B=48, while the Figure 23 on the right shows the computation.

It is again reasonable to assume that the number of filters is bounded, independently of L, while the number of temporal points depends on L [329]. As the role of M and N is switched in the assumption in Section 4.1.1 for the complexity, we arrive at a complexity of O (LlogL). This is also the complexity of the FFT of the whole signal. So the overall complexity of the frequency-dependent nonstationary Gabor transform is O (LlogL). The advantage of the method in terms of computational efficiency thus decreases as longer signals are considered. We note that at this point, since the windows used are band-limited, the current procedure is not suitable for real-time processing, despite its efficiency. The next step would be to process the incoming samples in a piecewise manner, using only a single family of frame elements for signals of arbitrary length. This entails working on finite, discrete parts of the given signal, thus considering the Fourier-transformed versions of vectors f·h∈CL, where h denotes some function of length L≪L [330-335]. This window, together with the frame elements, will be designed to minimize undesired effects that stem from the cutting of the signal. Details of this piecewise processing, as well as a proposed variable-Q transform, will be further discussed in a future contribution.

Downsizing the binding epitope to a minimal functional descriptor in 3D: The recognition of IL10-R1 by apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b involves 23 receptor residues [336]. The particularly discontinuous arrangement of these residues in sequence, and their quite scattered disposition in 3D space along ca. 34 Å, makes it quite challenging to assemble all their functionalities within a small molecule able to mimic the apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_bR1 interaction to apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b. Based on available structural data [337], residues apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b have been proposed as being the most relevant for protein-receptor recognition. Our energetic calculations (see materials and methods) confirmed those residues as being the major contributors to the energy of binding of apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b. In our approach, as strategy to ease the definition of key functionalities in the targeted binding site, we downsized the number of residues to be taken into account by a selection procedure based on:

i) Calculated residue binding energy contribution,

ii) Interfacial hydrogen bond network with apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b, and

iii) Proximity in space. Based on these criteria, residues apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b, which exhibited top binding energies, an extensive interfacial hydrogen bond network with apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b and were disposed in close

proximity in space (all within 12 Å), were considered as key functionalities for our seeding template search [338]. Though downsizing the large binding epitope to three residues would imply at first none or considerably low binding affinity of any seeding template molecule containing them, the geometric and functional constraints demarcated by the side chains of the selected residues would constitute the minimal 3D functional descriptors to generate regular expressions signifying relevant functionalities of the targeted recognition site to be mimicked (Figure 24, Table 10).

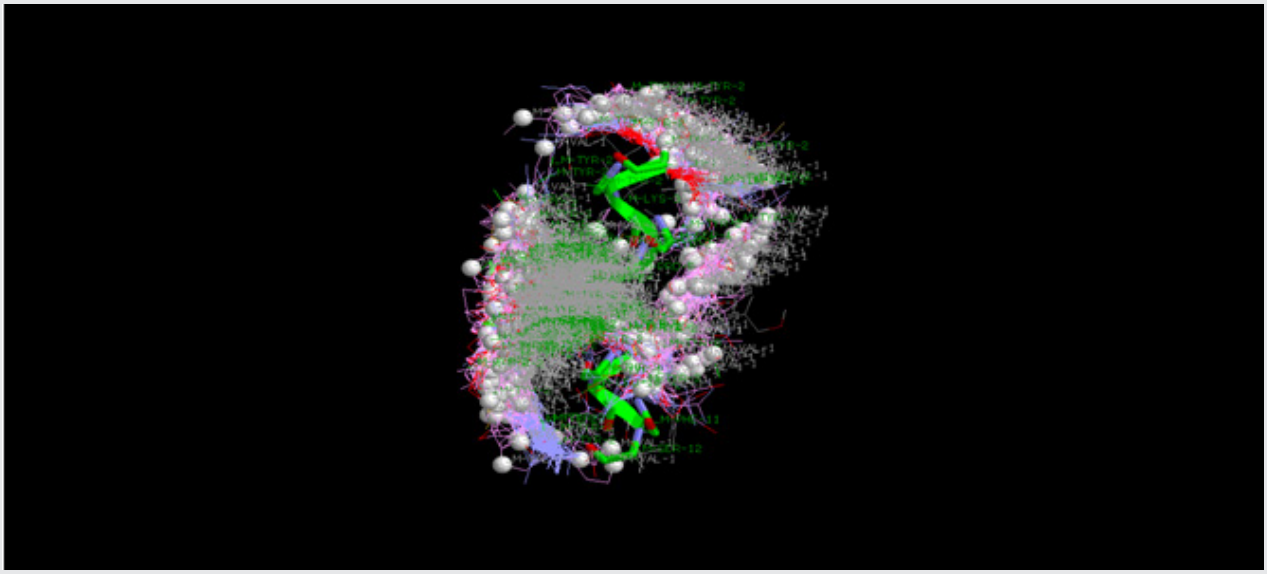


Figure 24

Table 10:

VYKSPNAYTLFS-ZINC35032631-0.pdb	-59.0033	-28.0159	-30.9874	0	18.9167
VYKSPNAYTLFS-ZINC35032482-0.pdb	-55.2906	-51.7906	-3.5	0	20.2941
VYKSPNAYTLFS-ZINC35031704-0.pdb	-53.6269	-35.3997	-18.2272	0	20.3571
VYKSPNAYTLFS-ZINC35031701-0.pdb	-50.4926	-47.9926	-2.5	0	18.4706
VYKSPNAYTLFS-ZINC35030592-1.pdb	-59.5908	-56.9333	-2.65749	0	19.6667
VYKSPNAYTLFS-ZINC35028323-0.pdb	-57.8179	-51.819	-5.9989	0	19.375
VYKSPNAYTLFS-ZINC35027939-0.pdb	-50.3806	-42.0499	-8.33076	0	18.5
VYKSPNAYTLFS-ZINC35016046-0.pdb	-56.4917	-39.1091	-17.7234	0.340666	19.4
VYKSPNAYTLFS-ZINC35016025-0.pdb	-54.0756	-44.5798	-9.4958	0	22.3333
VYKSPNAYTLFS-ZINC35016022-0.pdb	-55.3572	-46.4325	-8.9247	0	18.7333
VYKSPNAYTLFS-ZINC34951876-0.pdb	-62.4605	-43.1866	-19.6388	0.364929	18.4706
VYKSPNAYTLFS-ZINC34947823-1.pdb	-54.4351	-50.9351	-3.5	0	19.7333
VYKSPNAYTLFS-ZINC34926226-0.pdb	-58.7188	-43.2983	-15.4206	0	22.0769
VYKSPNAYTLFS-ZINC34924833-0.pdb	-55.2602	-45.1438	-10.1164	0	20.7143
VYKSPNAYTLFS-ZINC34924832-1.pdb	-54.8991	-43.0808	-11.8183	0	21.5
VYKSPNAYTLFS-ZINC34902561-1.pdb	-50.2396	-43.4677	-6.77188	0	20.2308
VYKSPNAYTLFS-ZINC34892606-1.pdb	-52.9569	-42.1431	-10.8138	0	18.5
VYKSPNAYTLFS-ZINC34887010-0.pdb	-57.1822	-57.1822	0	0	19
VYKSPNAYTLFS-ZINC34875592-1.pdb	-49.12	-39.62	-9.5	0	18.4615
VYKSPNAYTLFS-ZINC34875187-1.pdb	-58.7221	-46.4019	-12.3201	0	21.9375
VYKSPNAYTLFS-ZINC34875186-0.pdb	-55.5807	-41.7166	-13.8641	0	19.625
VYKSPNAYTLFS-ZINC34849893-1.pdb	-50.5269	-36.5398	-13.987	0	20.0769

VYKSPNAYTLFS-ZINC34788433-1.pdb	-45.0131	-36.8049	-8.20822	0	22.7273
VYKSPNAYTLFS-ZINC34788430-1.pdb	-48.2371	-35.3442	-12.8928	0	22.4545
VYKSPNAYTLFS-ZINC34788427-0.pdb	-50.0478	-35.4176	-14.6302	0	18.6429
VYKSPNAYTLFS-ZINC34788423-0.pdb	-49.3081	-46.8081	-2.5	0	19.9286
VYKSPNAYTLFS-ZINC34787057-1.pdb	-51.94	-41.8296	-10.1105	0	19.4
VYKSPNAYTLFS-ZINC34776496-0.pdb	-68.9384	-36.97	-31.9683	0	18.6923
VYKSPNAYTLFS-ZINC34770532-0.pdb	-58.2325	-44.2325	-14	0	20.6154
VYKSPNAYTLFS-ZINC34770531-0.pdb	-58.3567	-44.3567	-14	0	20.8462
VYKSPNAYTLFS-ZINC34770257-0.pdb	-51.1638	-34.536	-16.6278	0	20.2143
VYKSPNAYTLFS-ZINC34768560-0.pdb	-45.8149	-37.3189	-8.49596	0	23.6
VYKSPNAYTLFS-ZINC34763656-0.pdb	-58.3077	-44.3077	-14	0	19.3571
VYKSPNAYTLFS-ZINC34763652-0.pdb	-51.1137	-44.1137	-7	0	20.9286
VYKSPNAYTLFS-ZINC34761710-1.pdb	-52.8246	-45.8246	-7	0	21.3571
VYKSPNAYTLFS-ZINC34759726-0.pdb	-48.2778	-41.2782	-6.99956	0	21.5833
VYKSPNAYTLFS-ZINC34759722-0.pdb	-51.2653	-40.2732	-10.9921	0	19.6667
VYKSPNAYTLFS-ZINC34753843-1.pdb	-65.3117	-46.4922	-18.8195	0	19.8125
VYKSPNAYTLFS-ZINC34753842-1.pdb	-58.4898	-40.4654	-18.0244	0	18.5
VYKSPNAYTLFS-ZINC34753733-0.pdb	-49.7142	-38.3415	-11.3728	0	20.2308
VYKSPNAYTLFS-ZINC34753615-1.pdb	-45.7788	-26.1664	-19.6123	0	26.4444
VYKSPNAYTLFS-ZINC34753614-1.pdb	-50.4741	-26.7953	-23.992	0.313162	24.3333
VYKSPNAYTLFS-ZINC34753360-0.pdb	-48.233	-44.733	-3.5	0	20.3333
VYKSPNAYTLFS-ZINC34744441-1.pdb	-52.8383	-36.8496	-15.9887	0	19.2941
VYKSPNAYTLFS-ZINC34744439-0.pdb	-64.5915	-41.3653	-23.2262	0	18.5294
VYKSPNAYTLFS-ZINC34744437-0.pdb	-58.8762	-35.8049	-23.0713	0	18.4706
VYKSPNAYTLFS-ZINC34744435-1.pdb	-57.0243	-41.9218	-15.1025	0	18.6471
VYKSPNAYTLFS-ZINC34729638-0.pdb	-52.497	-43.9952	-8.50181	0	19
VYKSPNAYTLFS-ZINC34729635-0.pdb	-53.4082	-46.283	-7.12513	0	21
VYKSPNAYTLFS-ZINC34628799-0.pdb	-58.8135	-48.742	-10.0715	0	21.6429
VYKSPNAYTLFS-ZINC34628320-0.pdb	-50.9941	-38.1996	-13.1623	0.367857	20.3077
VYKSPNAYTLFS-ZINC34628319-1.pdb	-48.7856	-39.2856	-9.5	0	18.4615
VYKSPNAYTLFS-ZINC34627996-0.pdb	-39.1611	-32.1611	-7	0	22
VYKSPNAYTLFS-ZINC34627152-0.pdb	-55.3854	-45.9079	-9.47743	0	18.7059
VYKSPNAYTLFS-ZINC34626586-0.pdb	-50.5478	-47.0478	-3.5	0	18.625
VYKSPNAYTLFS-ZINC34626582-1.pdb	-53.7004	-47.7004	-6	0	20.0667
VYKSPNAYTLFS-ZINC34624614-1.pdb	-51.2203	-44.9342	-6.28609	0	19.9231
VYKSPNAYTLFS-ZINC34621787-0.pdb	-49.6183	-41.7989	-7.81944	0	19
VYKSPNAYTLFS-ZINC34617645-1.pdb	-51.5264	-35.3286	-16.4466	0.248732	23.1
VYKSPNAYTLFS-ZINC34616245-1.pdb	-54.6167	-51.1167	-3.5	0	19.1875
VYKSPNAYTLFS-ZINC34614593-1.pdb	-43.5229	-29.5229	-14	0	22.7143
VYKSPNAYTLFS-ZINC34613363-0.pdb	-48.6226	-43.6228	-4.99979	0	22.0833
VYKSPNAYTLFS-ZINC34611341-0.pdb	-51.0215	-36.4185	-14.603	0	18.5714

VYKSPNAYTLFS-ZINC34611340-1.pdb	-50.5536	-47.0665	-3.48712	0	20.9286
VYKSPNAYTLFS-ZINC34611339-0.pdb	-58.3178	-45.1636	-13.1542	0	23
VYKSPNAYTLFS-ZINC34611338-0.pdb	-57.0789	-40.4489	-16.63	0	18.5
VYKSPNAYTLFS-ZINC34611007-0.pdb	-54.846	-46.1104	-8.73566	0	19.875
VYKSPNAYTLFS-ZINC34601304-1.pdb	-54.4359	-40.4359	-14	0	20.7273
VYKSPNAYTLFS-ZINC34601303-1.pdb	-50.2116	-36.2239	-13.9877	0	21.2727
VYKSPNAYTLFS-ZINC34596136-0.pdb	-44.0042	-27.0042	-17	0	22.8571
VYKSPNAYTLFS-ZINC34592649-1.pdb	-43.766	-33.266	-10.5	0	20.3
VYKSPNAYTLFS-ZINC34590642-0.pdb	-57.2038	-39.3999	-17.8039	0	23.7273
VYKSPNAYTLFS-ZINC34579789-1.pdb	-62.9102	-46.1194	-16.7908	0	20.4118
VYKSPNAYTLFS-ZINC34579209-0.pdb	-45.0762	-25.7236	-19.3526	0	23.75
VYKSPNAYTLFS-ZINC34579208-1.pdb	-45.4169	-29.9432	-15.4736	0	24
VYKSPNAYTLFS-ZINC34579143-0.pdb	-48.6282	-28.5429	-20.0853	0	26
VYKSPNAYTLFS-ZINC34579142-1.pdb	-45.7017	-25.351	-20.3506	0	23.5
VYKSPNAYTLFS-ZINC34574889-1.pdb	-50.6556	-38.1462	-12.5093	0	18.5
VYKSPNAYTLFS-ZINC34574888-0.pdb	-60.9554	-46.9554	-14	0	18.5625
VYKSPNAYTLFS-ZINC34574680-1.pdb	-60.1131	-50.6453	-9.46779	0	20.1875
VYKSPNAYTLFS-ZINC34571646-0.pdb	-54.8258	-40.8309	-13.9949	0	19.3571
VYKSPNAYTLFS-ZINC34569717-0.pdb	-47.717	-38.217	-9.5	0	19.6364
VYKSPNAYTLFS-ZINC34567952-1.pdb	-56.8868	-47.1544	-9.73243	0	19.625
VYKSPNAYTLFS-ZINC34567947-1.pdb	-54.6815	-39.8485	-14.833	0	18.5294
VYKSPNAYTLFS-ZINC34567945-0.pdb	-51.0024	-42.4719	-8.5305	0	18.8125
VYKSPNAYTLFS-ZINC34567935-1.pdb	-59.8674	-40.0729	-19.7945	0	18.75
VYKSPNAYTLFS-ZINC34567932-0.pdb	-58.4604	-46.881	-11.5794	0	19.1765
VYKSPNAYTLFS-ZINC34567930-1.pdb	-47.7422	-30.0813	-18.0965	0.435573	18.4375
VYKSPNAYTLFS-ZINC34567927-1.pdb	-46.9936	-41.9991	-4.9945	0	18.625
VYKSPNAYTLFS-ZINC34567673-0.pdb	-52.0067	-43.6633	-8.34344	0	21.3077
VYKSPNAYTLFS-ZINC34567670-1.pdb	-53.5203	-43.9656	-9.55463	0	22.3846
VYKSPNAYTLFS-ZINC34567373-1.pdb	-58.0468	-34.1427	-23.9041	0	21.5833
VYKSPNAYTLFS-ZINC34567104-1.pdb	-53.4235	-32.5896	-20.834	0	20.5
VYKSPNAYTLFS-ZINC34561654-1.pdb	-51.7971	-33.0231	-18.774	0	19.8182
VYKSPNAYTLFS-ZINC34559924-0.pdb	-46.0665	-32.7299	-13.3367	0	21.2727
VYKSPNAYTLFS-ZINC34557241-1.pdb	-42.3828	-35.45	-6.93278	0	19
VYKSPNAYTLFS-ZINC34557240-0.pdb	-43.6055	-33.1055	-10.5	0	20.0833
VYKSPNAYTLFS-ZINC34557239-1.pdb	-40.8927	-33.9598	-6.93289	0	19
VYKSPNAYTLFS-ZINC34557238-0.pdb	-42.4653	-35.4653	-7	0	19.3636
VYKSPNAYTLFS-ZINC34552885-1.pdb	-45.3933	-38.4076	-6.98569	0	20.7273
VYKSPNAYTLFS-ZINC34552848-0.pdb	-57.3892	-44.5961	-12.7931	0	20.0667
VYKSPNAYTLFS-ZINC34552406-1.pdb	-54.7481	-45.2481	-9.5	0	21
VYKSPNAYTLFS-ZINC34551303-0.pdb	-54.48	-41.1148	-13.3652	0	20.1538
VYKSPNAYTLFS-ZINC34541600-0.pdb	-52.8983	-39.0032	-13.8951	0	18.6429

VYKSPNAYTLFS-ZINC34541048-1.pdb	-58.0057	-48.6838	-9.32193	0	19
VYKSPNAYTLFS-ZINC34541046-1.pdb	-60.3863	-46.3863	-14	0	18.5294
VYKSPNAYTLFS-ZINC34539694-1.pdb	-49.7451	-31.1403	-18.6049	0	22.1111
VYKSPNAYTLFS-ZINC34537793-1.pdb	-53.9113	-46.9113	-7	0	18.7333
VYKSPNAYTLFS-ZINC34537792-0.pdb	-51.6152	-42.1152	-9.5	0	18.6667
VYKSPNAYTLFS-ZINC34529374-0.pdb	-51.4012	-31.8079	-19.5933	0	22.4
VYKSPNAYTLFS-ZINC34529373-1.pdb	-50.4539	-24.5877	-25.8662	0	19
VYKSPNAYTLFS-ZINC34529002-1.pdb	-46.4895	-39.4895	-7	0	20.9091
VYKSPNAYTLFS-ZINC34529001-1.pdb	-43.3972	-37.6158	-5.78135	0	21.3636
VYKSPNAYTLFS-ZINC34527351-1.pdb	-64.4162	-41.7319	-22.6844	0	18.4444
VYKSPNAYTLFS-ZINC34516859-0.pdb	-60.1623	-48.0121	-12.1502	0	21.1429
VYKSPNAYTLFS-ZINC34514729-0.pdb	-45.7703	-38.7703	-7	0	22.9
VYKSPNAYTLFS-ZINC34514728-0.pdb	-45.9365	-38.9365	-7	0	24.2
VYKSPNAYTLFS-ZINC34514727-1.pdb	-44.509	-37.509	-7	0	23.7778
VYKSPNAYTLFS-ZINC34514726-0.pdb	-45.1049	-38.1094	-6.99556	0	23.3333
VYKSPNAYTLFS-ZINC34502526-0.pdb	-48.6625	-39.7033	-8.95924	0	18.5
VYKSPNAYTLFS-ZINC34500149-1.pdb	-63.6756	-51.129	-12.5467	0	19.3529
VYKSPNAYTLFS-ZINC34498487-0.pdb	-59.1756	-48.6374	-10.5383	0	20.4118
VYKSPNAYTLFS-ZINC34498401-1.pdb	-56.3614	-41.821	-14.5404	0	19.9231
VYKSPNAYTLFS-ZINC34498399-0.pdb	-57.851	-34.7941	-23.0569	0	18.5625
VYKSPNAYTLFS-ZINC34496883-0.pdb	-47.8991	-45.3991	-2.5	0	18.4706
VYKSPNAYTLFS-ZINC34492113-0.pdb	-52.7199	-32.3432	-20.6584	0.281662	18.5
VYKSPNAYTLFS-ZINC34492109-0.pdb	-53.8359	-40.9908	-12.8452	0	20.5714
VYKSPNAYTLFS-ZINC34492045-1.pdb	-57.4097	-48.7343	-8.67533	0	21.6667
VYKSPNAYTLFS-ZINC34490273-1.pdb	-62.4344	-50.4344	-12	0	19.1176
VYKSPNAYTLFS-ZINC34490069-1.pdb	-59.9896	-52.9896	-7	0	20.1765
VYKSPNAYTLFS-ZINC34490067-1.pdb	-59.2848	-52.6248	-6.66007	0	18.5556
VYKSPNAYTLFS-ZINC34490065-0.pdb	-63.3981	-56.3981	-7	0	21.0588
VYKSPNAYTLFS-ZINC34489224-0.pdb	-45.3769	-32.2978	-13.4567	0.377568	23.2222
VYKSPNAYTLFS-ZINC34479753-1.pdb	-52.4461	-45.4461	-7	0	20
VYKSPNAYTLFS-ZINC34477542-0.pdb	-62.7036	-52.2036	-10.5	0	20
VYKSPNAYTLFS-ZINC34477541-1.pdb	-50.2019	-34.4061	-16.0702	0.274371	20.5
VYKSPNAYTLFS-ZINC34477540-0.pdb	-47.5876	-36.367	-11.2206	0	22.7273
VYKSPNAYTLFS-ZINC34475831-0.pdb	-58.7455	-49.2455	-9.5	0	18.75
VYKSPNAYTLFS-ZINC34475830-0.pdb	-62.7724	-52.7126	-10.0598	0	20.5
VYKSPNAYTLFS-ZINC34475829-1.pdb	-59.5661	-50.0661	-9.5	0	19.1875
VYKSPNAYTLFS-ZINC34475828-1.pdb	-59.6321	-51.1405	-8.49168	0	20.3125
VYKSPNAYTLFS-ZINC34469114-1.pdb	-65.3235	-48.059	-17.2646	0	19.5556
VYKSPNAYTLFS-ZINC34464194-1.pdb	-49.7706	-36.9625	-12.8081	0	22.8
VYKSPNAYTLFS-ZINC34464193-1.pdb	-51.9414	-35.6523	-16.2891	0	25
VYKSPNAYTLFS-ZINC34464192-0.pdb	-47.6577	-33.5617	-14.0959	0	24.1

VYKSPNAYTLFS-ZINC34464191-1.pdb	-56.608	-26.3173	-30.6302	0.339485	21
VYKSPNAYTLFS-ZINC34462809-0.pdb	-46.4501	-25.2449	-21.2053	0	19.3333
VYKSPNAYTLFS-ZINC34454786-0.pdb	-44.8621	-32.3533	-12.8635	0.354718	24.1111
VYKSPNAYTLFS-ZINC34454785-1.pdb	-42.2837	-35.2837	-7	0	22.3333
VYKSPNAYTLFS-ZINC34452039-1.pdb	-57.6843	-40.7577	-16.9266	0	22.6667
VYKSPNAYTLFS-ZINC34433807-1.pdb	-51.7121	-38.2664	-13.4457	0	20.625
VYKSPNAYTLFS-ZINC34427725-1.pdb	-59.6897	-45.8133	-13.8765	0	19.7143
VYKSPNAYTLFS-ZINC34417532-1.pdb	-69.3807	-52.6936	-16.6871	0	19.5556
VYKSPNAYTLFS-ZINC34417531-1.pdb	-70.0355	-47.8721	-22.1633	0	18.9444
VYKSPNAYTLFS-ZINC34412143-1.pdb	-56.8141	-51.3822	-5.43189	0	18.8235
VYKSPNAYTLFS-ZINC34406717-1.pdb	-52.9359	-49.4359	-3.5	0	19
VYKSPNAYTLFS-ZINC34370211-0.pdb	-53.2251	-40.1289	-13.0961	0	18.75
VYKSPNAYTLFS-ZINC34370210-0.pdb	-54.5979	-40.0206	-14.5773	0	18.5
VYKSPNAYTLFS-ZINC34370209-1.pdb	-57.4226	-38.5039	-18.9187	0	19.1875
VYKSPNAYTLFS-ZINC34370208-1.pdb	-56.5261	-41.7385	-14.7876	0	18.5
VYKSPNAYTLFS-ZINC34329749-0.pdb	-56.8711	-38.5208	-18.3503	0	23.6364
VYKSPNAYTLFS-ZINC34329748-1.pdb	-65.4229	-33.093	-32.3299	0	19.6364
VYKSPNAYTLFS-ZINC34321021-0.pdb	-55.644	-45.1921	-10.4519	0	19.3333
VYKSPNAYTLFS-ZINC34321020-0.pdb	-51.5124	-41.0188	-10.4936	0	18.7333
VYKSPNAYTLFS-ZINC34321019-1.pdb	-57.0079	-46.6158	-10.3921	0	19.0625
VYKSPNAYTLFS-ZINC34321018-0.pdb	-51.8658	-41.3658	-10.5	0	18.5
VYKSPNAYTLFS-ZINC34304487-1.pdb	-51.5188	-33.0929	-18.4259	0	20.8182
VYKSPNAYTLFS-ZINC34288034-1.pdb	-42.167	-32.667	-9.5	0	25.1429
VYKSPNAYTLFS-ZINC34287807-0.pdb	-59.7291	-43.7037	-16.0253	0	18.7059
VYKSPNAYTLFS-ZINC34287806-0.pdb	-59.2658	-44.8873	-14.3786	0	19
VYKSPNAYTLFS-ZINC34287805-1.pdb	-59.7547	-37.3127	-22.6814	0.239372	18.4706
VYKSPNAYTLFS-ZINC34287804-0.pdb	-50.4024	-33.2785	-17.1239	0	18.4706
VYKSPNAYTLFS-ZINC34278228-0.pdb	-38.7919	-31.8072	-6.9848	0	25.1429
VYKSPNAYTLFS-ZINC34235675-0.pdb	-56.1387	-43.9834	-12.1552	0	21.7333
VYKSPNAYTLFS-ZINC34182444-0.pdb	-52.1436	-39.7402	-12.4034	0	18.8824
VYKSPNAYTLFS-ZINC34161425-1.pdb	-39.9679	-25.9679	-14	0	21.8571
VYKSPNAYTLFS-ZINC34160134-0.pdb	-45.693	-30.6738	-15.3941	0.374855	21.4444
VYKSPNAYTLFS-ZINC34160133-0.pdb	-45.3508	-31.6934	-14.0211	0.363644	20.5556
VYKSPNAYTLFS-ZINC34139950-1.pdb	-52.4027	-28.9016	-23.5011	0	23.2222
VYKSPNAYTLFS-ZINC34139949-1.pdb	-47.8277	-36.1509	-11.6768	0	23.5556
VYKSPNAYTLFS-ZINC34139948-1.pdb	-44.6049	-27.4852	-17.1197	0	21.4444
VYKSPNAYTLFS-ZINC34129110-1.pdb	-56.7498	-46.2532	-10.4967	0	21.2857
VYKSPNAYTLFS-ZINC34122494-0.pdb	-56.3628	-46.1463	-10.2165	0	19.5333
VYKSPNAYTLFS-ZINC34116219-0.pdb	-56.4088	-41.8374	-14.5713	0	18.5714
VYKSPNAYTLFS-ZINC34082726-0.pdb	-53.9616	-34.9616	-19	0	18.5882
VYKSPNAYTLFS-ZINC34076870-1.pdb	-64.7002	-51.7135	-12.9867	0	21.5333

VYKSPNAYTLFS-ZINC34076869-0.pdb	-57.5791	-48.0983	-9.48076	0	19.1333
VYKSPNAYTLFS-ZINC34075807-1.pdb	-56.5795	-33.2224	-23.3572	0	23
VYKSPNAYTLFS-ZINC34027303-0.pdb	-47.8933	-34.3132	-13.58	0	25.4444
VYKSPNAYTLFS-ZINC33947808-0.pdb	-59.9722	-49.9374	-10.0348	0	20.6667
VYKSPNAYTLFS-ZINC33939040-1.pdb	-60.3517	-34.6725	-25.6791	0	20.2727
VYKSPNAYTLFS-ZINC33765559-1.pdb	-63.7726	-51.7541	-12.0185	0	19.3333
VYKSPNAYTLFS-ZINC33739862-0.pdb	-46.9141	-43.4141	-3.5	0	20.5
VYKSPNAYTLFS-ZINC33693876-0.pdb	-49.0912	-42.1699	-6.92134	0	22.75
VYKSPNAYTLFS-ZINC33506585-1.pdb	-56.2692	-36.0123	-20.2568	0	20.4286
VYKSPNAYTLFS-ZINC33378647-0.pdb	-61.9392	-53.1482	-8.79091	0	20.8235
VYKSPNAYTLFS-ZINC33376898-1.pdb	-54.2697	-37.8273	-16.4423	0	22.8571
VYKSPNAYTLFS-ZINC33376445-1.pdb	-50.8112	-38.349	-12.4622	0	19.0833
VYKSPNAYTLFS-ZINC32689261-1.pdb	-50.6494	-43.6947	-6.95477	0	19.4667
VYKSPNAYTLFS-ZINC32603235-0.pdb	-51.6817	-32.7103	-18.9714	0	22.5556
VYKSPNAYTLFS-ZINC32222731-0.pdb	-49.6471	-43.4595	-6.18753	0	22.8333
VYKSPNAYTLFS-ZINC32181558-0.pdb	-35.0443	-28.0443	-7	0	25.3333
VYKSPNAYTLFS-ZINC32180809-0.pdb	-56.4687	-46.1673	-10.3014	0	20.1333
VYKSPNAYTLFS-ZINC32107279-1.pdb	-61.507	-45.5513	-15.9558	0	18.4706
VYKSPNAYTLFS-ZINC32107277-0.pdb	-64.8584	-50.8584	-14	0	19.1176
VYKSPNAYTLFS-ZINC32107226-0.pdb	-65.2239	-46.8035	-18.4204	0	21.4375
VYKSPNAYTLFS-ZINC32107224-0.pdb	-57.5354	-45.7228	-11.8126	0	19.0625
VYKSPNAYTLFS-ZINC32107130-0.pdb	-56.5202	-33.402	-23.1182	0	18.5
VYKSPNAYTLFS-ZINC31776901-1.pdb	-58.0356	-44.2085	-13.8271	0	23
VYKSPNAYTLFS-ZINC31517997-1.pdb	-61.7249	-47.7249	-14	0	18.5625
VYKSPNAYTLFS-ZINC29553629-0.pdb	-60.7952	-47.9351	-12.86	0	19
VYKSPNAYTLFS-ZINC29553625-0.pdb	-54.2368	-45.2793	-9.32326	0.365662	18.6471
VYKSPNAYTLFS-ZINC29546418-0.pdb	-39.8244	-32.8244	-7	0	22.5556
VYKSPNAYTLFS-ZINC29546415-1.pdb	-39.8699	-32.8699	-7	0	21.4444
VYKSPNAYTLFS-ZINC28826366-0.pdb	-58.9713	-56.9443	-2.02702	0	20.25
VYKSPNAYTLFS-ZINC28826350-1.pdb	-65.8386	-61.343	-4.49562	0	20.5882
VYKSPNAYTLFS-ZINC28826348-0.pdb	-58.3626	-49.2583	-9.35709	0.252819	18.7333
VYKSPNAYTLFS-ZINC28826345-1.pdb	-63.5876	-50.5955	-12.9922	0	19.2353
VYKSPNAYTLFS-ZINC28826344-0.pdb	-59.4598	-45.5682	-13.8916	0	20
VYKSPNAYTLFS-ZINC28826341-1.pdb	-56.6987	-47.8375	-8.86115	0	18.4706
VYKSPNAYTLFS-ZINC28826340-1.pdb	-55.7956	-52.4112	-3.38436	0	19.8667
VYKSPNAYTLFS-ZINC28759422-1.pdb	-53.1489	-45.4777	-7.67116	0	18.5
VYKSPNAYTLFS-ZINC28704829-1.pdb	-51.9942	-40.6594	-11.3348	0	20.5
VYKSPNAYTLFS-ZINC28100854-1.pdb	-61.2097	-34.51	-26.6997	0	18.625
VYKSPNAYTLFS-ZINC28100848-1.pdb	-63.1965	-45.6965	-17.5	0	18.5625
VYKSPNAYTLFS-ZINC28100843-1.pdb	-50.0768	-47.5768	-2.5	0	18.6875

VYKSPNAYTLFS-ZINC28100838-0.pdb	-50.1862	-37.3707	-12.8155	0	18.625
VYKSPNAYTLFS-ZINC27553251-1.pdb	-67.3421	-50.4213	-16.9208	0	19.9412
VYKSPNAYTLFS-ZINC27553248-0.pdb	-64.9099	-51.5704	-13.3395	0	19.4118
VYKSPNAYTLFS-ZINC27192423-0.pdb	-60.9165	-43.3704	-17.5461	0	18.75
VYKSPNAYTLFS-ZINC27187558-0.pdb	-57.1815	-42.0153	-15.1662	0	20
VYKSPNAYTLFS-ZINC27187553-0.pdb	-56.5711	-32.2236	-24.3475	0	18.5
VYKSPNAYTLFS-ZINC27185422-1.pdb	-60.0396	-45.0292	-15.0104	0	19.7333
VYKSPNAYTLFS-ZINC27185410-1.pdb	-55.0055	-36.8101	-18.1954	0	18.4667
VYKSPNAYTLFS-ZINC27076149-1.pdb	-61.314	-42.8193	-18.4947	0	18.5714
VYKSPNAYTLFS-ZINC26838668-1.pdb	-57.1133	-40.6868	-16.4266	0	20.25
VYKSPNAYTLFS-ZINC26838664-1.pdb	-59.997	-27.8735	-32.1235	0	19.75
VYKSPNAYTLFS-ZINC26829224-0.pdb	-59.6076	-41.8574	-17.7503	0	19.0667
VYKSPNAYTLFS-ZINC26829220-1.pdb	-59.3786	-38.4251	-20.9535	0	19.3333
VYKSPNAYTLFS-ZINC26672590-1.pdb	-65.8979	-51.0704	-14.8275	0	19.1111
VYKSPNAYTLFS-ZINC26672586-1.pdb	-67.6052	-48.2232	-19.382	0	19.5556
VYKSPNAYTLFS-ZINC26547916-0.pdb	-44.0852	-31.5297	-12.9111	0.355501	24.3333
VYKSPNAYTLFS-ZINC26547913-0.pdb	-45.3508	-31.3265	-14.4592	0.434894	19
VYKSPNAYTLFS-ZINC26547891-1.pdb	-62.4925	-42.5481	-19.9444	0	19.2857
VYKSPNAYTLFS-ZINC26547888-0.pdb	-60.1436	-35.6061	-24.5375	0	18.5333
VYKSPNAYTLFS-ZINC26546949-0.pdb	-47.7053	-40.7053	-7	0	18.4286
VYKSPNAYTLFS-ZINC26546943-1.pdb	-45.6143	-38.6143	-7	0	19
VYKSPNAYTLFS-ZINC26021496-1.pdb	-56.8729	-31.3557	-25.5171	0	19.3125
VYKSPNAYTLFS-ZINC26019861-0.pdb	-63.4081	-34.5044	-28.9038	0	19.2
VYKSPNAYTLFS-ZINC26012015-1.pdb	-60.2255	-46.3297	-13.8958	0	19.7059
VYKSPNAYTLFS-ZINC26012012-1.pdb	-61.4981	-39.6111	-21.887	0	18.5882
VYKSPNAYTLFS-ZINC26012005-1.pdb	-60.1715	-41.8705	-18.301	0	19.0588
VYKSPNAYTLFS-ZINC26011999-0.pdb	-59.9766	-38.5079	-21.4687	0	20.2353
VYKSPNAYTLFS-ZINC25721242-0.pdb	-57.4088	-41.9761	-15.4327	0	20.9286
VYKSPNAYTLFS-ZINC25627094-1.pdb	-59.6005	-48.0653	-11.5351	0	19.4706
VYKSPNAYTLFS-ZINC25626937-1.pdb	-64.2475	-47.768	-16.4795	0	18.7059
VYKSPNAYTLFS-ZINC22204334-0.pdb	-67.9142	-46.539	-21.3752	0	18.4706
VYKSPNAYTLFS-ZINC22204315-1.pdb	-65.1417	-46.3464	-18.7953	0	19.5294
VYKSPNAYTLFS-ZINC22203891-1.pdb	-62.9235	-54.2623	-8.66127	0	19.8235
VYKSPNAYTLFS-ZINC22198994-0.pdb	-56.8364	-28.0401	-28.7964	0	18.7143
VYKSPNAYTLFS-ZINC22130459-1.pdb	-66.4794	-50.8036	-15.6757	0	20.0556
VYKSPNAYTLFS-ZINC22130406-0.pdb	-58.7567	-47.6271	-11.1296	0	21.1333
VYKSPNAYTLFS-ZINC22114256-1.pdb	-51.1572	-43.0803	-8.07692	0	20.6154
VYKSPNAYTLFS-ZINC22064434-0.pdb	-59.515	-45.0644	-14.4507	0	18.8
VYKSPNAYTLFS-ZINC22064432-1.pdb	-60.3454	-46.3467	-13.9987	0	19.2
VYKSPNAYTLFS-ZINC22058744-1.pdb	-59.9075	-37.4412	-22.4662	0	18.5882
VYKSPNAYTLFS-ZINC22058740-0.pdb	-58.6471	-51.16	-7.48714	0	19.4118

VYKSPNAYTLFS-ZINC22058732-1.pdb	-60.4677	-44.6825	-15.7852	0	18.4667
VYKSPNAYTLFS-ZINC22056810-0.pdb	-52.3701	-34.7854	-17.5847	0	22.3636
VYKSPNAYTLFS-ZINC22049550-1.pdb	-54.8932	-40.3007	-14.5925	0	23.7273
VYKSPNAYTLFS-ZINC22049545-1.pdb	-52.2278	-38.1098	-14.118	0	20.4167
VYKSPNAYTLFS-ZINC22049482-1.pdb	-57.1642	-37.686	-19.4782	0	18.4706
VYKSPNAYTLFS-ZINC22049477-1.pdb	-57.2593	-37.3717	-19.8876	0	18.4706
VYKSPNAYTLFS-ZINC22049472-1.pdb	-59.5881	-45.9197	-13.6684	0	19.4706
VYKSPNAYTLFS-ZINC22049467-0.pdb	-57.5837	-39.6205	-17.9633	0	18.5882
VYKSPNAYTLFS-ZINC22049463-0.pdb	-62.0983	-44.7595	-17.3388	0	18.4706
VYKSPNAYTLFS-ZINC22049458-1.pdb	-57.8672	-44.4136	-13.4536	0	18.4706
VYKSPNAYTLFS-ZINC22049453-0.pdb	-55.8994	-40.9842	-14.9152	0	19.9412
VYKSPNAYTLFS-ZINC22049448-0.pdb	-57.842	-43.0461	-14.7959	0	20.2857
VYKSPNAYTLFS-ZINC22049444-1.pdb	-57.1273	-42.4123	-14.7149	0	20
VYKSPNAYTLFS-ZINC22049439-0.pdb	-53.5036	-41.2195	-12.6368	0.352678	18.9286
VYKSPNAYTLFS-ZINC22049416-0.pdb	-54.9318	-34.3585	-20.5733	0	21.1429
VYKSPNAYTLFS-ZINC22049412-0.pdb	-59.6842	-28.9907	-30.6936	0	19.5714
VYKSPNAYTLFS-ZINC22049403-0.pdb	-47.8125	-28.6341	-19.1784	0	25.875
VYKSPNAYTLFS-ZINC22049399-0.pdb	-53.5335	-28.1324	-25.4011	0	23.625
VYKSPNAYTLFS-ZINC22046554-0.pdb	-53.9482	-43.5389	-10.4093	0	18.4706
VYKSPNAYTLFS-ZINC22046550-1.pdb	-52.1847	-44.7578	-7.42684	0	19.3529
VYKSPNAYTLFS-ZINC22009511-0.pdb	-47.8225	-35.8296	-11.993	0	20.4545
VYKSPNAYTLFS-ZINC22002320-1.pdb	-58.8268	-46.0452	-12.7816	0	20.0667
VYKSPNAYTLFS-ZINC22002315-0.pdb	-50.4873	-38.2239	-12.2634	0	20
VYKSPNAYTLFS-ZINC22002311-1.pdb	-53.5317	-38.2134	-15.3184	0	18.9333
VYKSPNAYTLFS-ZINC22002306-0.pdb	-55.1233	-42.6433	-12.4799	0	19.8667
VYKSPNAYTLFS-ZINC21993116-1.pdb	-53.8751	-32.7426	-21.1325	0	21.8182
VYKSPNAYTLFS-ZINC21990362-1.pdb	-50.9287	-45.5609	-5.36783	0	19.0667
VYKSPNAYTLFS-ZINC21990359-1.pdb	-59.0912	-44.9025	-14.1887	0	19.5333
VYKSPNAYTLFS-ZINC21990356-0.pdb	-54.565	-52.065	-2.5	0	18.8235
VYKSPNAYTLFS-ZINC21990354-1.pdb	-58.4766	-52.4766	-6	0	18.4706
VYKSPNAYTLFS-ZINC21989403-1.pdb	-49.5877	-42.5877	-7	0	18.4667
VYKSPNAYTLFS-ZINC21989401-1.pdb	-49.7468	-42.7468	-7	0	18.4667
VYKSPNAYTLFS-ZINC21989398-1.pdb	-49.1666	-49.1666	0	0	18.6667
VYKSPNAYTLFS-ZINC21989396-1.pdb	-48.8228	-39.9819	-8.84095	0	18.5333
VYKSPNAYTLFS-ZINC21985602-0.pdb	-61.1329	-30.3022	-30.8307	0	19.0909
VYKSPNAYTLFS-ZINC21985597-1.pdb	-55.7793	-46.2793	-9.5	0	18.9286
VYKSPNAYTLFS-ZINC21985368-1.pdb	-46.8542	-45.7367	-1.11746	0	20.5
VYKSPNAYTLFS-ZINC21985366-1.pdb	-49.3807	-48.3366	-1.04408	0	21.2857
VYKSPNAYTLFS-ZINC20426942-1.pdb	-55.8461	-44.0007	-11.8455	0	19.7857
VYKSPNAYTLFS-ZINC20232286-0.pdb	-58.4186	-28.3646	-30.054	0	18.4615
VYKSPNAYTLFS-ZINC20232281-1.pdb	-64.0325	-34.3518	-29.6807	0	19.3077

VYKSPNAYTLFS-ZINC19907272-1.pdb	-51.2166	-44.2166	-7	0	19.3333
VYKSPNAYTLFS-ZINC19893242-0.pdb	-59.9221	-54.2077	-5.71446	0	19.2941
VYKSPNAYTLFS-ZINC19891524-1.pdb	-59.1218	-52.2758	-6.846	0	19.8824
VYKSPNAYTLFS-ZINC19872842-0.pdb	-48.0035	-39.5006	-8.5029	0	20.1667
VYKSPNAYTLFS-ZINC19872840-1.pdb	-49.6994	-43.5221	-6.17738	0	22.8333
VYKSPNAYTLFS-ZINC19872837-1.pdb	-57.0293	-51.9532	-5.07608	0	18.5
VYKSPNAYTLFS-ZINC17724845-0.pdb	-58.8854	-36.4528	-22.4326	0	18.7059
VYKSPNAYTLFS-ZINC17421447-1.pdb	-49.6549	-47.1549	-2.5	0	19
VYKSPNAYTLFS-ZINC17378346-0.pdb	-45.988	-28.8378	-17.1502	0	21.8
VYKSPNAYTLFS-ZINC17314801-1.pdb	-56.9317	-45.9317	-11	0	20.3571
VYKSPNAYTLFS-ZINC17300590-0.pdb	-57.9089	-41.2597	-16.6492	0	20.0769
VYKSPNAYTLFS-ZINC17288490-1.pdb	-54.8833	-45.3833	-9.5	0	21.7692
VYKSPNAYTLFS-ZINC17197391-0.pdb	-60.3834	-31.7644	-28.9382	0.319212	19
VYKSPNAYTLFS-ZINC17176129-0.pdb	-48.4032	-41.6407	-6.76253	0	19.5714
VYKSPNAYTLFS-ZINC17175346-0.pdb	-49.5346	-33.0346	-16.5	0	20
VYKSPNAYTLFS-ZINC17175261-1.pdb	-49.5511	-40.1477	-9.40337	0	24.1
VYKSPNAYTLFS-ZINC17173515-1.pdb	-54.6811	-34.4521	-20.5052	0.276203	23.8
VYKSPNAYTLFS-ZINC17164548-0.pdb	-57.6872	-51.9364	-5.75077	0	20.1667
VYKSPNAYTLFS-ZINC17161511-0.pdb	-49.6984	-37.0476	-12.6508	0	19.6923
VYKSPNAYTLFS-ZINC17108604-1.pdb	-57.3681	-39.5468	-17.8213	0	22.2727
VYKSPNAYTLFS-ZINC17064522-1.pdb	-57.9748	-50.9748	-7	0	19.6875
VYKSPNAYTLFS-ZINC17063817-0.pdb	-55.7344	-43.8036	-11.9309	0	20.0714
VYKSPNAYTLFS-ZINC17063815-0.pdb	-60.1013	-43.2365	-16.8648	0	21.0714
VYKSPNAYTLFS-ZINC17047258-1.pdb	-58.3489	-43.3784	-14.9705	0	19.6429
VYKSPNAYTLFS-ZINC17001182-1.pdb	-55.4565	-40.9074	-14.5491	0	21.8462
VYKSPNAYTLFS-ZINC16997856-1.pdb	-49.1275	-32.4968	-16.9452	0.314591	20.4545
VYKSPNAYTLFS-ZINC16982373-0.pdb	-65.1802	-56.8691	-8.31107	0	21.3889
VYKSPNAYTLFS-ZINC16968896-0.pdb	-55.1697	-50.0923	-5.07744	0	20.7857
VYKSPNAYTLFS-ZINC16955443-1.pdb	-55.8506	-41.8508	-13.9998	0	20.8462
VYKSPNAYTLFS-ZINC16954002-1.pdb	-43.4843	-39.9886	-3.49564	0	19.5385
VYKSPNAYTLFS-ZINC16952697-0.pdb	-46.2516	-42.7516	-3.5	0	18.4706
VYKSPNAYTLFS-ZINC16952695-1.pdb	-49.7854	-46.3098	-3.47561	0	19.5333
VYKSPNAYTLFS-ZINC16941446-0.pdb	-53.9915	-40.9138	-13.0777	0	19.3125
VYKSPNAYTLFS-ZINC16941442-1.pdb	-51.6702	-38.1111	-13.5591	0	21.3077
VYKSPNAYTLFS-ZINC16940978-1.pdb	-44.722	-25.1469	-19.5751	0	20.625
VYKSPNAYTLFS-ZINC16940076-1.pdb	-56.3418	-41.3814	-14.9604	0	20.6667
VYKSPNAYTLFS-ZINC16923303-0.pdb	-51.0758	-48.5812	-2.49463	0	18.8125
VYKSPNAYTLFS-ZINC16696935-1.pdb	-56.3467	-46.8853	-9.46141	0	20.0588
VYKSPNAYTLFS-ZINC16696932-0.pdb	-55.2488	-44.152	-11.0968	0	20.75
VYKSPNAYTLFS-ZINC16696930-0.pdb	-62.1698	-45.3272	-16.8426	0	19.5882
VYKSPNAYTLFS-ZINC16123805-1.pdb	-51.2624	-38.6561	-12.6064	0	20.9167

VYKSPNAYTLFS-ZINC15937364-0.pdb	-56.0892	-46.5892	-9.5	0	18.5
VYKSPNAYTLFS-ZINC15937228-1.pdb	-52.7566	-45.8227	-6.93396	0	18.7143
VYKSPNAYTLFS-ZINC15937227-0.pdb	-52.3191	-36.9014	-15.6965	0.27889	18.5
VYKSPNAYTLFS-ZINC15935031-1.pdb	-54.155	-42.8319	-11.3231	0	18.5625
VYKSPNAYTLFS-ZINC15922260-1.pdb	-55.2799	-45.8123	-9.46764	0	18.5294
VYKSPNAYTLFS-ZINC15894679-1.pdb	-61.4056	-43.9606	-17.445	0	20.8
VYKSPNAYTLFS-ZINC15785332-1.pdb	-52.3768	-45.3768	-7	0	18.5333
VYKSPNAYTLFS-ZINC15785118-0.pdb	-52.4424	-37.7134	-14.7289	0	21.5455
VYKSPNAYTLFS-ZINC15784788-1.pdb	-48.6639	-36.252	-12.4119	0	19.6923
VYKSPNAYTLFS-ZINC15784786-0.pdb	-44.626	-33.453	-11.173	0	22.5
VYKSPNAYTLFS-ZINC15784784-1.pdb	-46.8674	-40.776	-6.09139	0	23.2727
VYKSPNAYTLFS-ZINC15784782-1.pdb	-46.434	-40.5153	-5.91874	0	23.3636
VYKSPNAYTLFS-ZINC15784780-1.pdb	-47.2555	-41.2121	-6.04336	0	23.3636
VYKSPNAYTLFS-ZINC15784546-0.pdb	-58.2402	-52.5056	-5.73467	0	20.4375
VYKSPNAYTLFS-ZINC15784544-0.pdb	-55.6665	-46.2791	-9.38742	0	21.0714
VYKSPNAYTLFS-ZINC15784358-0.pdb	-51.9988	-34.3997	-17.5991	0	20.8
VYKSPNAYTLFS-ZINC15784356-1.pdb	-50.8377	-47.3377	-3.5	0	20.9231
VYKSPNAYTLFS-ZINC15784354-1.pdb	-49.6606	-43.421	-6.2396	0	22.6667
VYKSPNAYTLFS-ZINC15783176-1.pdb	-55.645	-48.7516	-6.89332	0	20.625
VYKSPNAYTLFS-ZINC15782452-0.pdb	-53.6382	-34.0165	-19.6217	0	22.4545
VYKSPNAYTLFS-ZINC15781858-1.pdb	-64.1145	-50.1472	-13.9673	0	22.1429
VYKSPNAYTLFS-ZINC15781121-0.pdb	-55.4475	-41.9497	-13.4978	0	20.4615
VYKSPNAYTLFS-ZINC15780430-1.pdb	-66.5006	-40.3675	-26.1331	0	18.6
VYKSPNAYTLFS-ZINC15780220-0.pdb	-64.5033	-53.3748	-11.1285	0	19.1765
VYKSPNAYTLFS-ZINC15780218-0.pdb	-56.7917	-47.4859	-9.3058	0	20.5
VYKSPNAYTLFS-ZINC15779471-1.pdb	-56.3374	-42.3478	-13.9896	0	20.4167
VYKSPNAYTLFS-ZINC15778754-1.pdb	-56.5915	-46.0934	-10.4981	0	19.6
VYKSPNAYTLFS-ZINC15778697-0.pdb	-63.4126	-45.515	-18.1577	0.260111	19.4
VYKSPNAYTLFS-ZINC15443135-0.pdb	-51.5021	-37.5021	-14	0	20.9091
VYKSPNAYTLFS-ZINC15443133-1.pdb	-48.8158	-34.8668	-13.949	0	21.3636
VYKSPNAYTLFS-ZINC15443126-0.pdb	-62.1473	-32.5383	-29.609	0	20.0909
VYKSPNAYTLFS-ZINC15440710-1.pdb	-47.9766	-40.9766	-7	0	19.6923
VYKSPNAYTLFS-ZINC15440708-1.pdb	-45.0148	-38.0352	-6.97963	0	19.0769
VYKSPNAYTLFS-ZINC15022795-0.pdb	-53.7064	-40.5723	-13.134	0	23.0769
VYKSPNAYTLFS-ZINC15022783-0.pdb	-54.0612	-43.571	-10.4902	0	18.6

Template-based rescaffolding strategy

3D Regular expressions describing minimal 3D functional descriptors: The task of searching through repositories of protein structures is considerably more difficult than searching through protein sequence databases. The overall amount of flexibility in protein structures varies considerably, with certain substructures remaining relatively rigid, while others undergo considerable natural variation. Specifying a motif as a specific spatial arrangement of residues has certain weaknesses: firstly, in the computational difficulties of identifying matches, and secondly, in allowing enough flexibility in the pattern so that it can match a broad and meaningful set of structures. The first issue has been addressed in a number of studies such as the template library TESS by using geometric hashing techniques [339] or Jess, which builds on the previous approach by using a back tracking algorithm [340]. Flexibility in the

geometric constraints has been previously included as an overall threshold in the accepted RMSD between template and matches structures [341-347]. Our approach to the protein structure search problem has been to develop a procedure that allows us to define structural relationships between flexible sequence patterns. For this, we have taken the commonly used sequence-based PROSITE syntax [331,332] as bases to formulate a 3D pattern syntax, which allows a minimalist representation of relationships in 3D space among functionalities without the necessity of introducing any template coordinates as input file (see materials and methods for details).

The above-mentioned 3D functional descriptors and the geometric relationships among them has been used to generate the 3D syntax query R-<2,6>-R-<4,5>-Y-, which incorporates the selected key functionalities distributed in 3D as desired. Functionalities (side chains) are specified in our search interface as pseudo-points [348]. In the 3D syntax query, R-<2,6>-R describes a range of distances (2 to 6 Å) covering all possible spatial relationships among the side chains of apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b (being R76a and R76b the two conformers found for residue R76 in equal occupancy in the crystal structure). The relation between the functionalities of the side chains of apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b, or between one conformer of apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b are described by R-<4,5>-Y. Finally, the hyphen character "-"at the end of the R-<2,6>-R-<4,5>-Y- query specifies a cyclic pattern that reinforces tyrosine functionalities in a way that should be placed in space between the two arginine functionalities already described by the previous syntax elements.

Search and identification of seeding templates

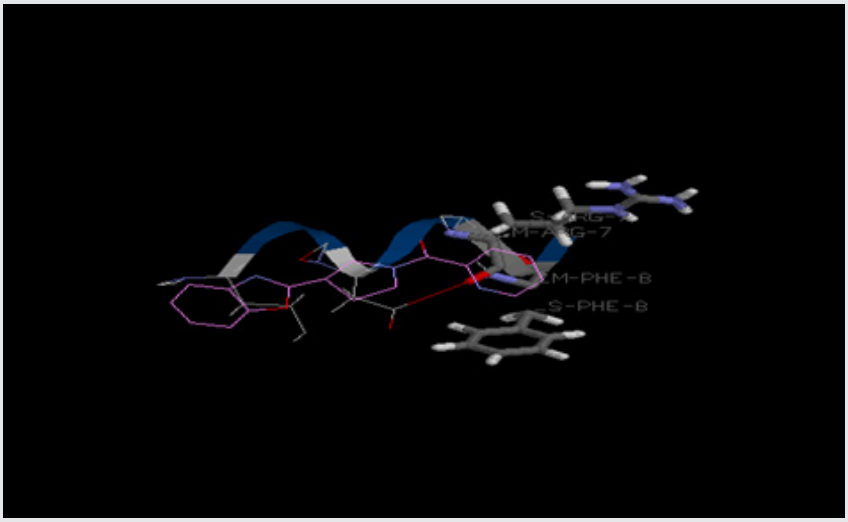


Figure 25

Our approach uses a 3D pattern query and a search algorithm to scrutinize the PDB. The millions of atomic coordinates collected in the PDB (ca. 115.000 experimentally obtained macromolecular structures) and organized in different 3D topologies accommodating endless combinations of functionalities are exploited to identify small regular architectures matching the selected 3D functional descriptors (see materials and methods for details). The scrutiny of the PDB apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b by our search engine with the query R-<2,6>-R-<4,5>-Y- threw off a total of 102 matches to our 3D functional descriptors. Only those hits containing the desired descriptors in a well-defined or regular secondary structure were considered. Furthermore, we selected the shortest architectures (up to 10 residues long) and, by rejecting those that were redundant, we were left with seven candidate structural motifs Table 1. Interestingly, these contained helical structures, indicating that a helix-like architecture could be considered as a suitable scaffold to achieve an appropriate 3D disposition of the required functionalities. Alpha-helical scaffolds are very attractive from the design point of view as they can be easily stabilized by chemical modifications and, furthermore, they offer many possibilities for substitutions [316,318,319]. The selected seven structural motifs were manually superimposed on the X-ray structure of the apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b complex such that a maximal overlap of their functionalities with the selected apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b binding residues was achieved. Those motifs with best atomic overlapping, 2ARZ105-109, 1ZYL127-131 and 2ACA151-155, were selected. These superposition models were used to analyze which motif would provide the possibility of introducing additional functionalities such that they would engage in the maximum number of complementary interactions with apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b. In this line, we adopted a straightforward strategy considering the possibility to elongate the alpha helix at the N- and C-termini [349]. 1ZYL127-131 and 2ACA151-155 were disposed longitudinally with respect to the protein-receptor interface, whereas 2ARZ105-109 was transversally oriented. The longitudinal orientation was able to provide a larger number of interactions with the protein when considering the N- and C- terminal elongation of the helical scaffold. Therefore, the 1ZYL and 2ACA structural motifs were taken into account with two and five additional residues at the N- and C-termini, respectively, which were already in helical conformation in their respective PDB structure. The resulting dodecameric helical scaffold X1X2R3Y4X5X6R7X8X9X10X11X12 (being

X any amino acid and underlined positions those representing the 3D functional descriptors) was considered as seeding template for further design purposes. Its model in complex with apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b was used to investigate in atomic detail those positions that would allow structure-stabilizing chemical modifications excluding positions 3, 4 and 7, which resemble the relevant functionalities of apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b for recognition (Figure 25).

Residues in apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b PDB matching the 3D pattern syntax query R-<2,6>-R-<4,5>-Y-.

Ligand Rational design strategy and experimental validation on conserved (β2GPI) antibodies, Anti-phospholipid antibodies (aPL) or from (PAPP-A; gi: 38045915) and APOH apolipoprotein H post-trancripts domains. Structure-based design of first generation apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_bR1 mimetics. Peptides of less than 15 residues derived from sequences present in protein helical domains usually lack such conformation when they are taken out of the stabilizing environment of the protein. Several approaches have been developed to stabilize short peptides into alpha helical conformations. The formation of covalent linkages between adjacent residues has been one of the preferred methodologies [311-313,316,318,319]. In this regard, lactam bridges between the side chains of lysine and aspartic acid in positions i,i + 4 have been shown to stabilize most efficiently short synthetic peptides into an alpha-helical structure [348]. We therefore adopted a lactam bridge design strategy. Our 12-mer seeding template X1X2R3Y4X5X6R7X8X9X10X11X12 (vide supra) offers the possibility to accommodate two i,i + 4 lactam bridges, which, in order to fix the desired conformation of positions 3, 4 and 7 with respect to apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b (i.e. in two consecutive helical turns), could be introduced separately or consecutively: VYKSPNAYTLFS and [K1X2R3Y4D5][K6R7X8X9D10]X11X12, respectively (square brackets represent lactam bridges between lysine (K) and aspartic acid (D) side chains). In the first scenario, the twelve residues would be constrained in helical conformation. In the second, only residues 1 to 10 would be constrained. Nevertheless, the helical conformation could be extended beyond these constraints at least by one or two more residues [350]. Based on our 3D molecular models, the first option would offer only position 11 for introducing additional H-bond interactions with apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b, whereas in the second option positions 8 and 11 would be available. In addition, our models suggested that the region of apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b interacting with the residue in position 11 (loop AB) would require certain flexibility in its counterpart. Therefore, two consecutive i,i + 4 lactam bridges were introduced in the scaffold resulting in a bicyclic molecule, which was then subjected to a per residue structure-based mutagenesis process in order to achieve best binding complementarity to apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b. Here, substitutions at positions 3, 4, 7, 8 and 11 pointed towards apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b, whereas residues at positions 9 and 12 did not establish contacts with the protein, and they were therefore maintained unaltered as in the corresponding PDB template structure (2ACA149-160). Position 12 could be used to enhance helicity in the C-term. Following this rationale, a lysine residue was introduced at position 8 to favor interactions with apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_bD44, and position 11 was used to introduce an arginine side chain for interactions with apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b. In order to analyze the possibility that the four positive charges so-far introduced in our bicyclic template could engage in non-specific electrostatic interactions at other regions of apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b, we performed energy interaction calculations with a sp2amine NH2 cation and a sp3 amine NH3 cation chemical probes by using GRID [351]. The most favorable energy was obtained at the predicted binding region for our template, suggesting that the selected topology of charged residues would confer sufficient binding specificity to the designed molecule. At this first stage, our rational design strategy also included additional strategic residue modifications directed towards the experimental evaluation of ligand binding at the protein targeted site by means of fluorescence spectroscopy (vide infra). A tryptophan or its 5-hydroxy derivative was introduced at position 2 of our ligands as labels to enable monitoring of protein complex formation. Furthermore, in order to enhance the sensitivity of tryptophan fluorescence measurements upon ligand binding at the targeted site, our rationale also made use of a mutation of cysteine to tyrosine in the protein binding region (i.e. at position 149, see below and Figure 26 for details). Thus, we could anticipate that, in case of ligand binding to the protein in the targeted site, the tryptophan emission produced would be enhanced by an energy transfer from the neighbor protein residue Y149. Finally, our template was acetylated and amidated in its N- and C-termini, respectively. The resulting apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_bR1 mimetic molecule Ac-[KW2R3Y4D][KR7K8VD]R11A-NH2 (M1, Table 2) was refined in complex with apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b using MD simulations (see materials and methods for details). The mimetic-protein binding free energy (ΔGM-P) was estimated with the MM-PBSA method and indicated favorable electrostatic and van der Waals contributions to the binding [352,353] Table 2. Hydrogen bonding between both molecules was evaluated. In the refined model, R3 participated in a hydrogen bond with apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b with apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b. K8 and R11 formed hydrogen bonds with apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b. Furthermore, the aliphatic moieties of the R3, Y4 and R7 side chains of the mimetic showed important van der Waals contributions with the side chains of apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b. The relevance of mimetic residues Y4, R5 and R7 in binding to apolipoprotein H [Homo sapiens] and

pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b was further investigated by MM-PBSA alanine mutagenesis [352-354]. The substitution of these residues by alanine revealed a considerable unfavorable effect in binding for all mutants, showing the most dramatic effect for alanine mutation at position 7 (R7A) (Figure 26).

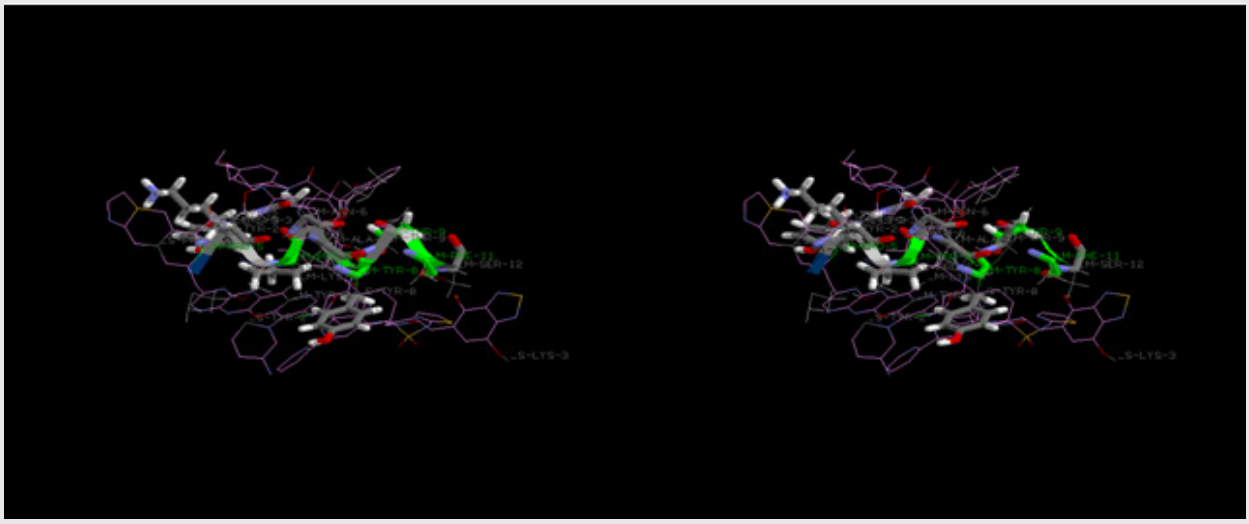


Figure 26

Structure-based design rationale (Figure 27)

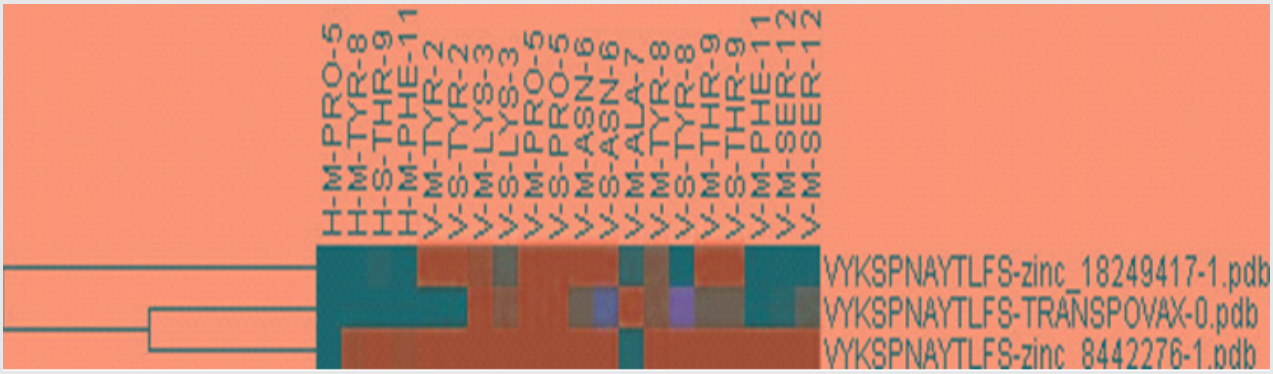


Figure 27

Mimetic-protein binding free energy obtained by MM-PBSA [352,353] and experimental dissociation constants (Kd). [a],[b]

Experimental validation of first generation apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_bR1 mimetics

M1.

Discussion and Conclusion

Reaction-diffusion systems for spatio-temporal intracellular protein networks: A beginner's guide with two examples4. In paper we have raised the question of replacing commonly used ODE models for conserved (β2GPI) antibodies, Anti-phospholipid antibodies (aPL) or from (PAPP-A; gi:38045915) and APOH apolipoprotein H post-trancripts domains protein-protein interactions by reaction-diffusion models which are well studied [30,31,43] and which not only contain reaction terms coming from the protein interactions but also describe the spatial distribution of species involved in the reactions over the cellular compartments. We have proposed the simpler Leloup-Goldbeter circadian model for (β2GPI) antibodies, Anti-phospholipid antibodies (aPL) in conserved (β2GPI) antibodies, Anti-phospholipid antibodies (aPL) or from (PAPP-A; gi: 38045915) and APOH apolipoprotein H post-trancripts domains that contains two equations for the nuclear and cytoplasmic (β2GPI) antibodies, Anti-phospholipid antibodies (aPL) protein and its mRNA, and a more complicated model for p53 response to the occurrence of DSBs in a single cell [355]. These models, and more generally reaction-diffusion PDE models for intracellular protein dynamics, are likely to be more realistic than ODE models. As an example of such possibly more realistic models, let us consider the slightly more complicated PER model in the conserved (β2GPI) antibodies, Anti-phospholipid antibodies (aPL) or from (PAPP-A; gi: 38045915) and APOH apolipoprotein H post-trancripts domains of Leloup and Goldbeter [44,45], that relies on the same principles as the (β2GPI) antibodies, Anti-phospholipid antibodies (aPL) model [356-358]. Indeed, two additional phosphorylations are considered, and this first PER model (more sophisticated models of PER have been published by Leloup and Goldbeter) is proposed to be amenable to describe two conserved (β2GPI) antibodies, Anti-phospholipid antibodies (aPL)

or from (PAPP-A; gi:38045915) and APOH apolipoprotein H post-trancripts domains mutants for PER, with shorter or longer period of oscillations, by lower and higher values, respectively, of the maximum degradation rate vd of cytoplasmic diphosphorylated PER [359]. But is this sole change of constants the best reason, in a more physiological model, for the change in periods? Could it be related to transcription of genes, to species diffusion, to translation into protein products, to cytosolic/nucleic degradation, to nucleocytoplasmic transport, to the extension of the "dead zone" around the nucleus where translation occurs? The same questions can be posed when one considers p53 and drugs amenable to electively modify specific parts of its intracellular dynamics [360-363]. This can be rendered optimally in PDE models, that are much more physiological than (rougher) ODE models, since they are naturally able to take intracellular spatial features into account, including possible space heterogeneities in the intracellular medium (not considered in the examples, apart from the above mentioned dead zone of non-translation in the cytoplasm). The oscillatory dynamics of conserved (β2GPI) antibodies, Anti-phospholipid antibodies (aPL) or from (PAPP-A; gi:38045915) and APOH apolipoprotein H post-trancripts domains, "the guardian of the genome", has long been evidenced by biological recordings in experimental conditions representing cell stress due to, e.g. radiotoxic insult [364-367]. In this sense, the biological question of identifying intracellular spatio-temporal dynamics is certainlsy crucial. Furthermore, since p53 disruptions are found in more than 50% of solid tumours, with various modifications of its stress-induced dynamics [144,145,132], it is also a crucial point in cancer therapy modelling to understand how conserved (β2GPI) antibodies, Anti-phospholipid antibodies (aPL) or from (PAPP-A; gi:38045915) and APOH apolipoprotein H dynamics may be affected in disease, and to this aim, to have an accurate representation of how this dynamics works in physiological conditions, which is the purpose of designing conserved (β2GPI) antibodies, Anti-phospholipid antibodies (aPL) or from (PAPP-A; gi:38045915) and APOH apolipoprotein H post-trancript dynamic models [368-372]. We have shown how easy it is in principle to start from an ODE compartmental model, add diffusion terms-which in their simplest version are mere laplacians and slightly modify the representation of exchanges between the compartments to adapt them to the new setting, to obtain a ready-made reaction-diffusion PDE model [373-376]. However, dependence on spatial patterns, such as oscillations in concentration, on the diffusivities and the (nuclear membrane) permeabilities has to be better studied, with precise identification of the underlying biological parameters, to understand their roles properly [377-380]. To obtain more realistic PDEs one can consider effects of different viscosities of buffers on different diffusivities of molecules or adjunction of advection terms when knowledge of aided transport makes it relevant (e.g., transport through microtubules). To support these ideas, different mobility of molecules in the nucleus and the cytoplasm caused by limited diffusions has led to biphasic caspase activation kinetics, see [381-384] and citations therein. In some circumstances it might be convenient to include specific spatial structures, such as position and density of ribosomes and thus to model gene expression more specifically [385,386]. In addition, intracellular signalling in our rather simplified modelling setting is restricted to cells of physiology and morphology where a molecular network in question is, at least, partially understood. What does the p53 signal transduction in response to DNA damage in a nerve cell or in a polynucleic muscle cell look like? What is the role of diffusion in a cell where a signal is spread over long distances and in very small cells? These and other questions should be further addressed and spatial PDE models might fruitfully be used for this purpose. Note also that the oscillatory patterns are self-organised not only due to events occurring in the nucleus or in the cytoplasm but that they are also tightly connected to boundary conditions on the membranes. Thus physiological delays maintained due to the semipermeability of the nuclear membrane are more typical for PDEs than for ODEs (if one does not want to deal with artificial delays represented by delay differential equations). We examined the Kedem-Katchalsky as representing passive transport mechanisms with the difference of concentrations at both sides of the membrane as the driving force for exchanges; however, bigger species are rather transported actively, which should be taken into account in more sophisticated models of nucleocytoplasmic transport [47]. On a more technical note, we have also briefly introduced the semi-implicit Rothe method that can be used for numerical solution of reaction-diffusion systems as an alternative to other used methods; they can be used also in 3D simulations, [37], where other methods are in general more demanding in time and memory. We hope that the simple presentation of this alternate solution to classical ODEs will be of some help to biologists and modellers who want to describe intracellular spatio-temporal dynamics of proteins in a faithful, yet more demanding in terms of parameter estimation, way. We would as well like to mention that such reaction-diffusion PDE models are also amenable to describe spatio-temporal dynamics at the level of cell populations and that, by introducing intercellular signalling, it is in principle possible to connect the two observation levels. This perspective still remains a challenge to mathematicians and modellers in biology. In this work, several 1,7-diazacarbazole analogs were identified as potentially effective oral Chk1 inhibitors through a series of computer-aided drug design processes, such as 3D-QSAR modeling, molecular docking, and molecular dynamics simulations. The CoMFA/CoMSIA models with statistical capacity showed good internal and external validation abilities and can be used to predict new and potential molecules. Moreover, the obtained contour maps can be used to guide the design of new compounds with high Chk1 inhibitory activity. Meanwhile, molecular docking and molecular dynamics process were established to study the possible binding modes of inhibitors at the active pocket of Chk1. Some key residues, such as Glu85, Cys87, Lys38, and Gly90, were found. Hydrogen bonding and electrostatic forces were predicted to be the key interactions that confer bioactivity. Overall, these results show that the optimal CoMFA/CoMSIA models can be used to predict novel Chk1 inhibitors and guide the development of new potential oral analogs.

Conclusion: The modern in silico (computational) chemical analysis respecting the bio- activity and availability of analogues substances, potentially beneficial or detrimental for specific interaction in organs and organisms, faces with a paradoxical dichotomy: if searching for the best correlation useful for prediction of specific molecular bio- or eco- activity QSAR models involving un-interpretable many latent variables may be produced, while always remaining the question of correlation factor indeterminacy (i.e. the assumed descriptors can be at any time replaced with other producing at least the same correlation performances); instead, when

restricting the analysis to search for molecular design and mechanisms throughout performing SARs by means of special structural indicators for a given class of relevant molecules, arises the price of limiting the use of generated models for further prediction. The present communication is mainly devoted in developing the second (Q)SAR facet by extending the recent introduced notion of spectral-path-linking-endpoints and the associate least action principle to spectral path quantification, in terms of the best fitted molecules, along the contained computed models, by means of the introduced q(uantum)-SAR factor within the generally called Quantum-SAR (QuaSAR) methodology. As an application, for representative flavonoids' inhibiting activities on breast cancer resistant protein there was clearly shown that the newly introduced q-SAR factor offers relevant analytical characterization of previously conceptually introduced spectral path hierarchy; moreover, the present QuaSAR may allow interpretation inter-conversion of concerned molecules' towards receptor binding since belonging to the same class of analogs, while they certainly undertaking such transformation during their interaction with macromolecules, proteins and enzymes present on cellular walls or with *in vivo* environment. Basically, the QuaSAR stands as the first step in assessing the quantum mechanically equivalent of wavefunction to the sample of molecules interacting with a specific organism site; it will eventually lead with the hyper-wave function with the help of which the associate hyperdensity probability of binding (metabolization) is to be computed; the last information may provide the density probability map of the ligand-receptor interaction abstractedfrom the structural Spectral-Qua-SAR correlations; with this tool the molecular design of new chemical structures may be appropriately undertaken. However, the actualQuaSAR scheme and quantum factor carry the main features of quantum dynamical systems and may stimulate future computational and conceptual developments inmolecular design for structurally controlled activity. Further generalization of the present QuaSAR method to modeling all potential inter-conversions of employedmolecules involved in correlation as well as for establishing their quantum metabolization complete map (through, for instance, hydrophobic, electrostatic and steric barriertunneling) is actually in progress and will be reported in subsequent communications. Our study of the influence of the amino acid reference energy shows the best energyfunction requires both the corrections for the dependency of energy terms and for individual amino acids. With both, the prediction accuracy surpasses 80% whenconsidering rotamers that are within top 6% of the lowest energy conformations. The over-counting of the energy can also be reverted by simply removing problematicenergy terms, suggesting that highly correlated energy calculations should be avoided when creating an energy model. However, neither the cross terms nor the amino acid reference energy could completely correct the false prediction biases towards some amino acids, which produces non-native like amino acid compositions in designedproteins. The distribution of predicted amino acids is comparable to the distribution of all proteins found in NCBI database only when distribution itself is introduced into the objective function, again showing the importance of a carefully constructed objective function. Finally, a concern was raised that the design test was too easy. We only required that the energy function to select the correct rotamer for a single position, leaving all neighboring side chains fixed. This is unrealistic in the sense that a real design problem would have several neighboring side chains unfixed. But this was the test we selected for two reasons. First, updating the energy for a single side chain after changing the parameters of the energy function would require repacking teh side chains. If this were done, the whole optimization process, which took tens of thousands of CPU hours as it is, would have taken millions of CPU hours. As such it was impossible even for the CCNI cluster. Second, we believed it was important to solve the easy problem first. If the energy function cannot predict the correct side chain given fixed neighbors, then there is no hope that it could solve the problem when the neighbors were not fixed. Others have sought to optimize the energy function by machine learning, each with a different slant. Yanover [385] used an objective function similar to our ΣlnP and trained a weighted sum of term (similar to Eq. 1) on the problem of side chain prediction. In Yanover's case, the identity of the amino acid was given. Even so, the prediction accuracy for the best method was only 82.6%, using a fine-grained rotamer library. As in our case, Ser and other polar side chains were the most problematic. Sharabi [386] optimized a linear sum of terms against a database of interface side chains. In this case, the energy terms that were up-weighted by the learning procedure made sense in the context of protein-protein interfaces. VDW was downweighted, as was the polar side chain burial (similar to our Eq. 7). Sharabi reasoned that since explicit water was not being considered, polar burial must be allowed. By optimizing for the recovery of side chain conformation, they were able to also recover the native sequence, which mirrors our result in which sequence-based training recovered the side chain structure. In both cases, the sequence answer was right for the right structural reason. Leaver-Fay [387] used a variety of structural metrics to evaluate three specific hypotheses to improve the Rosetta energy function. One metric was over/under prediction numbers, which we also used. This led to improvements in sequence recovery by smoothing the knowledge-based ϕ,ψ potential, a change that could not have been discovered by machine learning. Rosetta's sequence recovery rates (~38%) are better than ours (25%) when the energy is expressed as a linear sum (Eq. 1), but we were able to recover as much as 53% of the sequence when we added cross-terms with reference weights Table 5. We do not expect to develop an ideal energy function that eliminates all false designs, but having the ability in an energy function to accurately reduce the enormous space of all possible sequences down to a small set of suboptimal sequences that solve the design problem, makes feasible a faster and easier library screening in the wet lab. DG based solvation models have had a considerable success in solvation analysis 10,11,13,70. Particularly, our DG based non-polar solvation model was shown to offer some of the most accurate solvation energy predictions of various non-polar molecules.16 However, our DG based full solvation model is subject to numerical instability in solving the GLB equation, due to its coupling with the GPB equation. To stabilize the coupled GLB and GPB equations, a strong constraint on the van der Waals interaction was applied in our earlier work, 10,11,13 which hinders the parameter optimization of our DG based solvation model. In the present work, we resolve this problem by introducing new parameter optimization algorithms, namely, perturbation method and convex optimization, for the DG based solvation model. New stability conditions are explicitly imposed to the parameter selection, which guarantees the stability and robustness of solving the GLB equation and leads to constrained optimization of our DG based solvation model. The new optimization algorithms are

intensively validated by using a large number of test molecules, including the SAMPL0 test set, 53 alkane, alkene, ether, alcohol, and phenol types of solutes. Regression results based on our new algorithms are consistent extremely well with experimental data. Additionally, a fivefold cross validation technique is employed to explore the ability of our DG based solvation models for the blind prediction of the solvation free energies for a variety of solute molecules. It is found that the same level of errors is found in the training and validation sets, which confirms our model's predictive power in solvation free energy analysis. The present DG based full solvation model provides a unified framework for analyzing both polar and nonpolar molecules. In our future work, we will develop machine learning approaches for the robust clustering of solute molecules of interest into appropriate categories so as to better predict their solvation free energies This work demonstrated that the combination of PSO and SVMs can be applied to effectively and efficiently select major features in QSPR modeling of the thermodynamic parameters of 1:1 inclusion complexation of enantiomeric pairs of chiral guests with b-CD. This responds to the needs of drug designers for prediction of the thermodynamic parameters of new compounds in complexation with b-CD. The method was based on a discrete binary modification of PSO. The fitness function was the Int. J. Mol. Sci. 2009, 10 2119 Pearson correlation which was curve fitted by SVMs. The modified PSO appeared to be an effective and efficient algorithm, which robustly finds near-optimal and consistent results with short computer code and simple mathematical operators, while converging rather quickly. The SVMs showed excellent performance in predicting ln K, ΔG°, ΔH° and TΔS°, by considering major selected features. The combination of the adopted methods showed satisfactory results with the large dataset. We can use Shannon entropy measures to developing predictive models for multi-target networks of neuroprotective/neurotoxic compounds. In doing so, we can use Box-Jenkins operators of molecular descriptors to obtain multi-target, multi-scale, and multi-output models able to predict different outcomes for multiple combinations of output experimental measures, experimental protocols, organisms, and molecular and cellular targets. One of these models has been demonstrated here to be useful as a complementary tool in the organic synthesis and evaluation of the multi-target biological activity of new compounds with potential neuroprotective activity. The model is also a very useful tool to predict complex networks of drug-target interactions with possible applications to the study of non-linear effects in the biological activity of neuroprotective drugs. In recent years, attention has been paid to QSAR/QSPR methods as an interesting complement, or even as an expensive, time consuming alternative,

to laboratory data. In this paper, new QSPR models have been developed for predicting the tR of a diverse set of mycotoxins from the molecular structure alone. We have compared two linear models, MLR and SVM, with the data set. The obtained results show that both MLR and SVM methods could model the relationship between tR and their electronic and thermodynamic descriptors; on the same sets of descriptors, using SVM based produced a better model with a better predictive ability than the MLR model. SVM exhibit the better overall performance due to embodying the structural risk minimization principle and some advantages over the other techniques of converging to the global optimum and not to a local optimum. By performing model validation, it can be concluded that the presented model is a valid model and can be effectively used to predict the tR of mycotoxins with an accuracy approximating the accuracy of experimental tR determination. Moreover, the mechanism of the model was interpreted, and the applicability domain of the model was defined. It can be reasonably concluded that the proposed model would be expected to predict tR for new organic compounds or for other organic compounds for which experimental values are unknown. Additionally, the presented method could also identify and provide some insight into what structural features are related to the tR property of organic compounds. Simcyp is an ADME simulator with the ability to include inter-individual variability within DDI prediction [70,72,78]. Previously, the utility of Simcyp to predict DDIs has been investigated across a range of inhibitors and victim drugs, also assessing the impact of the time course of inhibition and induction [82]. However, the majority of comparisons of DDI prediction between the static and dynamic model used the static model equation (4), or the static model in Simcyp compared with an alternative dynamic model [271] rather than both models within Simcyp. Therefore, the parameters utilized have not been consistent, which confounds the direct comparison between the two models. In addition, some of the studies focus on only one inhibitor [123,272,273] or one victim drug [113,274] and conclusions may therefore be specific to that drug without the ability to ascertain more general trends. In contrast, the current study performs predictions for both the dynamic and static models in Simcyp, allowing consistency in parameters and direct assessment of the impact of the time course. This study comprehensively assessed the prediction accuracy of reversible DDIs of azole inhibitors and different victim drugs, including a representative number of weak, medium and strong interactions. Despite a plethora of research available with these combinations there is still no consensus on certain practices. In addition, a sufficient number of studies per inhibitor or victim drug allowed identification of trends and individual assessment of parameters in the current study. The impact of different parameters including active metabolites of inhibitor, dose timing and ka was investigated, as well as the ability of Simcyp to predict inter-individual variability in the extent of DDIs. The ability of Simcyp to incorporate inter-individual variability and to generate a range of output values gives an advantage over the equation-based static model (equation 1). It also allows a successful level of prediction when using the unbound drug concentrations, therefore conforming to the free drug hypothesis. This is in contrast to some previous publications advocating the use of total concentration based on the prediction success and number of studies estimated within the 2-fold margin [227,232]. In the current analysis, the Simcyp static model gave consistently higher results than the dynamic model, as the ratio of the predicted AUC ratios by dynamic: static model were predominantly <1. This is not surprising considering the consistently higher average inhibitor concentration used in static predictions. A comparable number of studies were found within 2-fold of the observed AUC ratio regardless of the model used, with similar bias (afe of 1.68 and 1.34) and precision (rmse of 3.81 and 3.61) for the dynamic or static model, respectively. Use of the weighting according to the number of subjects in the clinical studies did not affect bias and precision of predictions performed in the dynamic model (afe and rmse of 1.64 and 3.61, respectively). Any benefit from incorporating the time-

course of the inhibitor in the dynamic model was relatively minor for ketoconazole despite its short half-life (approximately 3.3±1h). The largest differences between static and dynamic model predictions were observed for interactions involving itraconazole or triazolam. The difference observed with itraconazole studies was likely to be explained by the inclusion of the active metabolite, where studies with a longer time scale included the additional action of hydroxy-itraconazole. The pie charts indicate that the data are not skewed by a disproportionate amount of DDI studies involving any particular inhibitor or victim drug. Good overall prediction of DDIs investigated was observed using both the dynamic and static models with 71 and 77% of studies within 2-fold of the observed AUC ratio, respectively. However, a trend for under-prediction was observed (using the parameter input values detailed in Table 2 rather than Simcyp default values) particularly involving the strong interactions (with ketoconazole and itraconazole), regardless of the model used. Five studies (12 data points) were included in the subset where DDIs were reported after single dosing of the inhibitor. One of the assumptions behind the prediction model shown in equation-1 is that the inhibitor concentrations are at steady-state. As demonstrated by the simulations performed by Zhao et al. [241], this would result in significant differences in the DDI magnitude for substrates with a long t1/2. The single dose studies in this subset only involve midazolam or triazolam which both have relatively short half-lives. Predictions from virtual trials for these victim drugs support the findings by Zhao et al. [241] and were not affected by the single dosing regimen of the inhibitor. However, the limited number of clinical studies in this range confounds any firm conclusions. No clear trend was observed to support a difference between static and dynamic DDI predictions or the accuracy against the observed clinical AUC ratio in terms of the bioavailability (0.32-0.88 for midazolam and alprazolam, respectively) [275] or t1/2 (1.5-12 h for triazolam and alprazolam, respectively) of the victim drugs in the dataset. This lack of trend was still apparent when other victim drugs were included in the assessment (i.e. simvastatin and ciclosporin, t1/2 of 2 and 6h, respectively, and bioavailability (F) <0.3). The inclusion of the most abundant and potent active metabolite of itraconazole resulted in improved prediction accuracy in the dynamic model, demonstrating the importance of this parameter in DDI predictions Figure 3. Simcyp currently allows the inclusion of one metabolite into DDI prediction, excluding the other two potentially relevant itraconazole active metabolites, keto- and N-desalkyl-itraconazole. N-desalkyl-itraconazole has a particularly long t1/2 and a comparable ratio of steady-state average unbound concentration and Ki to the hydroxy metabolite [89]. Therefore, it may contribute to the persistence of inhibition after itraconazole administration. The variation in itraconazole ka of approximately 2-fold of the clinical range in the dynamic model resulted in marginal difference in predicted AUC ratio (<11%). The impact may be seen for the larger range in ka (e.g. at 10-fold as simulated in Zhao et al. [123], but assuming that the range reported in the clinical study was proportional to the overall patient population, this 10 fold range is unlikely to be observed. Good prediction accuracy was observed for DDIs involving alprazolam or midazolam as the victim drug, with the majority of these DDIs predicted within 2-fold of the observed value. However, there was an under-prediction trend observed for triazolam DDIs regardless of the model used. Therefore, the impact of alternative in vitro clearance and permeability data used as input parameters was assessed. Minimal improvement in AUC ratio prediction accuracy was seen when triazolam in-house permeability data (from either Caco-2 or MDCK-MDR1 cells) [266] were used in the Qgut model to estimate triazolam FG values, which were consequently used in DDI prediction. The triazolam CLint utilized in AUC ratio predictions represented a mean value from four different studies performed in human liver microsomes. However, in one of these studies [134], a 20-fold difference in CLint was observed across the six donors, resulting in a maximum 5-fold difference (range of 2.20 to 9.54) in AUC ratio in the Simcyp simulations. This AUC ratio difference is not observed in individual trials, but can be observed across the whole virtual population.

This difference could be due to other confounding factors rather than a direct result of the CLint variability. Differences in predicted AUC ratio resulting from the use of CLint from human liver microsomes and the default recombinant values in Simcyp were minor in the case of triazolam. In the current study impact of different dosing regimens (staggered dosing) of the inhibitor on DDI prediction was assessed using an itraconazole interaction. In addition to midazolam (previously studied in a ketoconazole-midazolam interaction pair [123], triazolam and alprazolam were also included in the assessment (Table 11) (Figure 28).

Investigation found that the most pronounced difference based solely on virtual trials was observed with midazolam, and the maximum AUC ratio was observed with simultaneous dosing. However, the same trends were not observed when compared with the observed clinical data with different dosing schedules. For example, the maximum AUC ratio in one study with triazolam [120] was actually observed with substrate dosing 3 h after the inhibitor dose. However, this observed difference may be a result of the study design, considering that food intake was included with the inhibitor dose in this scenario and not when the inhibitor and substrate were simultaneously dosed. The decreases in the AUC for the dosing time >3 h for both midazolam and triazolam was over-estimated in the simulations in comparison with the observed AUC ratios [120]. However, limited availability of clinical data for different dosing regimens and the wide range of DDI magnitude observed for the same dosing regimen confounds any clear trends. The differential impact of dosing time between previous work with ketoconazole [123] and current analysis with itraconazole cannot be explained by the approximately 4-fold longer t1/2 of itraconazole (21±6h) in comparison with ketoconazole [276]. The prediction of inter-individual variability in DDI magnitude was assessed in comparison with the one study where AUC values of the victim drug ± inhibitor were reported for each individual [122]. In the case of ketoconazole, the ratio of median predicted: observed values across the 10 virtual trials ranged from 0.29-0.52 highlighting the under-prediction trend. However, predicted inter-individual variability in the AUC ratio reflected observed variability, as the coefficient of variation across the trials (45%) was consistent with the actual study (50%). A similar under-prediction trend was observed in the case of itraconazole, but to a lesser extent, as the ratio of median predicted/observed values was 0.74. However, variability estimated across trials for this interaction pair (32% coefficient of variation) was considerably lower in comparison with the observed data with this inhibitor (105%), associated with one significant outlier. If this subject was excluded from the analysis of the itraconazole DDI with triazolam, the predicted variability and differences between minimum and maximum AUC ratio are comparable to the observed data.

Assessment of different physiological, demographic and genetic characteristics of the individual subjects in the virtual population was performed in order to rationalize the variability seen in virtual subjects' AUC ratio, although corresponding data from *in vivo* trials were not available for direct comparison. This included simulated CYP3A4 content, based on a meta-analysis of liver data [235] or intestinal data from 31 individuals [277], which could only partially explain the inter-individual variability in AUC ratios. The prediction success of DDIs can only be validated against the reported *in vivo* values. However, it is advisable to consider these data very critically, considering that variability can be introduced as a result of inconsistency in study designs and data analysis. The small size of the population observed in this DDI study (nine subjects) or overall in the studies investigated (mean 10.0 ± 2.76, range 4-20 subjects) may result in a reduced ability to predict accurately population values, either by limiting the variability observed, or increasing the power of one subject if a significantly different result is observed, as illustrated in the case of itraconazole-triazolam. Overall, this study has reinforced the value of Simcyp as a successful platform to assess the prediction of DDIs in a population based model, as high prediction accuracy was observed for 35 DDIs. An under-prediction trend was observed in the current dataset for strong DDIs, with no overall significant difference in the prediction accuracy between the static and dynamic model. This lack of major difference between the two models can be attributed to the fact that concentrations of azoles are substantially higher in comparison with their Ki regardless of the approach used. However, this cannot be extrapolated to less potent or inhibitors administered at a low dose. Prediction accuracy was related to the particular victim drug in the DDI, with high prediction accuracy observed for midazolam DDIs and lower prediction accuracy seen for triazolam DDIs. Incorporation of the time course of inhibition was mainly beneficial for itraconazole, due to the ability to include the contribution of the active metabolite in the prediction model. An increasing number of compounds with multiple interaction mechanisms or contributing inhibitory metabolites [83,85,113] emphasize the importance of the dynamic model approach. Predicted inter-individual variability in DDI magnitude reflected the variability observed *in vivo* despite some under-prediction of mean, and highlighted the importance of critical evaluation of both the input parameters used in the simulations and the clinical data used for validation of the prediction success. In this work, we made use of the MM-PBSA technique in docking scoring and in affinity prediction of protein-peptide complexes. We also compared the results with the HADDOCK built-in scoring function. Overall, HADDOCK and dMM-PBSA, a dampened MM-PBSA version, behaved similarly in ranking the near-native poses in the top 3 clusters, improving over the standard MM-PBSA version. The introduction of weights for the different MM-PBSA terms is not unprecedented in the literature (Zhou et al., 2009), but this approach has never been applied to PPIs. Notably, despite the fact that different experimental conditions (where the main difference regarded the type of buffer and the ionic strength, whereas as both pH and temperature were comparable) and techniques Table1 were used to determine the dissociation constants, we observed a good correlation between experimental and computational ΔG of binding using HADDOCK and dMM-PBSA scoring functions, 0.63 and 0.66 respectively. Interestingly, lack of modulation of the solvation MM-PBSA terms resulted in worse correlation between the experimental and simulated Figures (r = 0.49). Our findings were then validated on two additional systems, one with known structure and binding affinity and one for which only the ΔGbind has been reported. Both HADDOCK and dMM-PBSA methods perform remarkably well in ranking the two additional protein-peptide complexes, and lead to good correlation with experimentally measured Δgbind. In order to assess the presence of possible systematic errors in the binding free energy calculations, we used two different PB solvers, namely APBS (Baker et al., 2001) and DelPhi (Rocchia et al., 2001, 2002). No differences in term of correlation with experimental data were found using the two different solvers, except that in absolute ΔGcomp values equation 1 and in the calculation speed (see Section Materials and Methods). APBS provided a ΔGcomp larger than DelPhi. This could arise from different aspects, including the different approaches used to describe the dielectric interface, the approach used to estimate the reaction field energy, and/or how the different solvers treat cavities that are internal to the solute. The accuracy of the PB equation solution has been reported to be sensitive to the grid size, in favor of smaller grid spacing (Sørensen et al., 2015). However, decreasing the grid spacing increases the computational resources needed to perform the calculation, both in terms of physical memory, and the computational time required. We choose a grid size of 0.5Å for both solvers since it represents a good trade-off between speed and accuracy. With lower grid resolution, APBS would have encountered problems in convergence in ΔGcomp calculation Sørensen et al., 2015. Finally, the choice of the interior dielectric (εint) value in the PB calculation is not trivial since in the literature its value can be found spanning between 1 and 20. Higher εint (4-20) aims to effectively mimic polarization and local rearrangement effects, as well as transient penetration of water molecules into the solute interior. They should in principle be preferred when the PB calculation is performed on individual structures. On the other hand, lower εint (2-3) is preferred in order to mainly account for electronic polarization and it is commonly used on ensembles of structures, which explicitly account for conformational flexibility. In light of this, there is still no consensus on the most appropriate εint value. We decided to use εint = 2 in all cases, while being aware that scaling the polar contribution in dMM-PBSA is similar to considering a dielectric screening of the solute medium, accounting for polarization and rearrangement due to the reaction to the existing fields. This is also in agreement with previous studies where the rank-ordering performance for MM-PBSA improves with increasing dielectric constant (Wang et al., 2013). Simultaneously, this supports the relative importance of the non-polar components and results in better performance. In fact, the van der Waals and BSA terms, which are directly related to the MM-PBSA non-polar terms, correlated well with the experimental data and they provided the main contribution to the final score. These factors might explain the better performance of the dampened MM-PBSA. Finally, we analyzed the calculated data from HADDOCK and MM-PBSA in order to evaluate the reproducibility of the results in terms of scoring and final correlation with the experimental data. First of all, for a given complex, with the corresponding MM energy terms provided by HADDOCK, the only variations in the final results could arise from the PB calculations since it depends on many parameters, including the grid spacing, the atomic radii, and the dielectric interior. For this reason, we used two different PB solvers, APBS and DelPhi, using identical parameters. In Figures 1,2 it is shown the good agreement in the MM-PBSA and dMM-PBSA calculations with the two PB solvers, indicating that the calculations are quite robust. Second, we calculated the correlations between the experimental and calculated

binding free energy using different MM-PBSA values from the clusterBEST4 and not the average value. The final correlations R were in a range of 0.39-0.49 for MM-PBSA and 0.60-0.67 for dMM-PBSA, indicating that in principle the MM-PBSA could be performed on a single pose with less computational effort. Finally, the calculated standard error for each clusterBEST4 reported in Table 11.

Table 11:

VYKSPNAYTLFS-ZINC35032631-0.pdb	-59.0033	-28.0159	-30.9874	0	18.9167
VYKSPNAYTLFS-ZINC35032482-0.pdb	-55.2906	-51.7906	-3.5	0	20.2941
VYKSPNAYTLFS-ZINC35031704-0.pdb	-53.6269	-35.3997	-18.2272	0	20.3571
VYKSPNAYTLFS-ZINC35031701-0.pdb	-50.4926	-47.9926	-2.5	0	18.4706
VYKSPNAYTLFS-ZINC35030592-1.pdb	-59.5908	-56.9333	-2.65749	0	19.6667
VYKSPNAYTLFS-ZINC35028323-0.pdb	-57.8179	-51.819	-5.9989	0	19.375
VYKSPNAYTLFS-ZINC35027939-0.pdb	-50.3806	-42.0499	-8.33076	0	18.5
VYKSPNAYTLFS-ZINC35016046-0.pdb	-56.4917	-39.1091	-17.7234	0.340666	19.4
VYKSPNAYTLFS-ZINC35016025-0.pdb	-54.0756	-44.5798	-9.4958	0	22.3333
VYKSPNAYTLFS-ZINC35016022-0.pdb	-55.3572	-46.4325	-8.9247	0	18.7333
VYKSPNAYTLFS-ZINC34951876-0.pdb	-62.4605	-43.1866	-19.6388	0.364929	18.4706
VYKSPNAYTLFS-ZINC34947823-1.pdb	-54.4351	-50.9351	-3.5	0	19.7333
VYKSPNAYTLFS-ZINC34926226-0.pdb	-58.7188	-43.2983	-15.4206	0	22.0769
VYKSPNAYTLFS-ZINC34924833-0.pdb	-55.2602	-45.1438	-10.1164	0	20.7143
VYKSPNAYTLFS-ZINC34924832-1.pdb	-54.8991	-43.0808	-11.8183	0	21.5
VYKSPNAYTLFS-ZINC34902561-1.pdb	-50.2396	-43.4677	-6.77188	0	20.2308
VYKSPNAYTLFS-ZINC34892606-1.pdb	-52.9569	-42.1431	-10.8138	0	18.5
VYKSPNAYTLFS-ZINC34887010-0.pdb	-57.1822	-57.1822	0	0	19
VYKSPNAYTLFS-ZINC34875592-1.pdb	-49.12	-39.62	-9.5	0	18.4615
VYKSPNAYTLFS-ZINC34875187-1.pdb	-58.7221	-46.4019	-12.3201	0	21.9375
VYKSPNAYTLFS-ZINC34875186-0.pdb	-55.5807	-41.7166	-13.8641	0	19.625
VYKSPNAYTLFS-ZINC34849893-1.pdb	-50.5269	-36.5398	-13.987	0	20.0769
VYKSPNAYTLFS-ZINC34788433-1.pdb	-45.0131	-36.8049	-8.20822	0	22.7273
VYKSPNAYTLFS-ZINC34788430-1.pdb	-48.2371	-35.3442	-12.8928	0	22.4545
VYKSPNAYTLFS-ZINC34788427-0.pdb	-50.0478	-35.4176	-14.6302	0	18.6429
VYKSPNAYTLFS-ZINC34788423-0.pdb	-49.3081	-46.8081	-2.5	0	19.9286
VYKSPNAYTLFS-ZINC34787057-1.pdb	-51.94	-41.8296	-10.1105	0	19.4
VYKSPNAYTLFS-ZINC34776496-0.pdb	-68.9384	-36.97	-31.9683	0	18.6923
VYKSPNAYTLFS-ZINC34770532-0.pdb	-58.2325	-44.2325	-14	0	20.6154
VYKSPNAYTLFS-ZINC34770531-0.pdb	-58.3567	-44.3567	-14	0	20.8462
VYKSPNAYTLFS-ZINC34770257-0.pdb	-51.1638	-34.536	-16.6278	0	20.2143
VYKSPNAYTLFS-ZINC34768560-0.pdb	-45.8149	-37.3189	-8.49596	0	23.6
VYKSPNAYTLFS-ZINC34763656-0.pdb	-58.3077	-44.3077	-14	0	19.3571
VYKSPNAYTLFS-ZINC34763652-0.pdb	-51.1137	-44.1137	-7	0	20.9286
VYKSPNAYTLFS-ZINC34761710-1.pdb	-52.8246	-45.8246	-7	0	21.3571
VYKSPNAYTLFS-ZINC34759726-0.pdb	-48.2778	-41.2782	-6.99956	0	21.5833
VYKSPNAYTLFS-ZINC34759722-0.pdb	-51.2653	-40.2732	-10.9921	0	19.6667

VYKSPNAYTLFS-ZINC34753843-1.pdb	-65.3117	-46.4922	-18.8195	0	19.8125
VYKSPNAYTLFS-ZINC34753842-1.pdb	-58.4898	-40.4654	-18.0244	0	18.5
VYKSPNAYTLFS-ZINC34753733-0.pdb	-49.7142	-38.3415	-11.3728	0	20.2308
VYKSPNAYTLFS-ZINC34753615-1.pdb	-45.7788	-26.1664	-19.6123	0	26.4444
VYKSPNAYTLFS-ZINC34753614-1.pdb	-50.4741	-26.7953	-23.992	0.313162	24.3333
VYKSPNAYTLFS-ZINC34753360-0.pdb	-48.233	-44.733	-3.5	0	20.3333
VYKSPNAYTLFS-ZINC34744441-1.pdb	-52.8383	-36.8496	-15.9887	0	19.2941
VYKSPNAYTLFS-ZINC34744439-0.pdb	-64.5915	-41.3653	-23.2262	0	18.5294
VYKSPNAYTLFS-ZINC34744437-0.pdb	-58.8762	-35.8049	-23.0713	0	18.4706
VYKSPNAYTLFS-ZINC34744435-1.pdb	-57.0243	-41.9218	-15.1025	0	18.6471
VYKSPNAYTLFS-ZINC34729638-0.pdb	-52.497	-43.9952	-8.50181	0	19
VYKSPNAYTLFS-ZINC34729635-0.pdb	-53.4082	-46.283	-7.12513	0	21
VYKSPNAYTLFS-ZINC34628799-0.pdb	-58.8135	-48.742	-10.0715	0	21.6429
VYKSPNAYTLFS-ZINC34628320-0.pdb	-50.9941	-38.1996	-13.1623	0.367857	20.3077
VYKSPNAYTLFS-ZINC34628319-1.pdb	-48.7856	-39.2856	-9.5	0	18.4615
VYKSPNAYTLFS-ZINC34627996-0.pdb	-39.1611	-32.1611	-7	0	22
VYKSPNAYTLFS-ZINC34627152-0.pdb	-55.3854	-45.9079	-9.47743	0	18.7059
VYKSPNAYTLFS-ZINC34626586-0.pdb	-50.5478	-47.0478	-3.5	0	18.625
VYKSPNAYTLFS-ZINC34626582-1.pdb	-53.7004	-47.7004	-6	0	20.0667
VYKSPNAYTLFS-ZINC34624614-1.pdb	-51.2203	-44.9342	-6.28609	0	19.9231
VYKSPNAYTLFS-ZINC34621787-0.pdb	-49.6183	-41.7989	-7.81944	0	19
VYKSPNAYTLFS-ZINC34617645-1.pdb	-51.5264	-35.3286	-16.4466	0.248732	23.1
VYKSPNAYTLFS-ZINC34616245-1.pdb	-54.6167	-51.1167	-3.5	0	19.1875
VYKSPNAYTLFS-ZINC34614593-1.pdb	-43.5229	-29.5229	-14	0	22.7143
VYKSPNAYTLFS-ZINC34613363-0.pdb	-48.6226	-43.6228	-4.99979	0	22.0833
VYKSPNAYTLFS-ZINC34611341-0.pdb	-51.0215	-36.4185	-14.603	0	18.5714
VYKSPNAYTLFS-ZINC34611340-1.pdb	-50.5536	-47.0665	-3.48712	0	20.9286
VYKSPNAYTLFS-ZINC34611339-0.pdb	-58.3178	-45.1636	-13.1542	0	23
VYKSPNAYTLFS-ZINC34611338-0.pdb	-57.0789	-40.4489	-16.63	0	18.5
VYKSPNAYTLFS-ZINC34611007-0.pdb	-54.846	-46.1104	-8.73566	0	19.875
VYKSPNAYTLFS-ZINC34601304-1.pdb	-54.4359	-40.4359	-14	0	20.7273
VYKSPNAYTLFS-ZINC34601303-1.pdb	-50.2116	-36.2239	-13.9877	0	21.2727
VYKSPNAYTLFS-ZINC34596136-0.pdb	-44.0042	-27.0042	-17	0	22.8571
VYKSPNAYTLFS-ZINC34592649-1.pdb	-43.766	-33.266	-10.5	0	20.3
VYKSPNAYTLFS-ZINC34590642-0.pdb	-57.2038	-39.3999	-17.8039	0	23.7273
VYKSPNAYTLFS-ZINC34579789-1.pdb	-62.9102	-46.1194	-16.7908	0	20.4118
VYKSPNAYTLFS-ZINC34579209-0.pdb	-45.0762	-25.7236	-19.3526	0	23.75
VYKSPNAYTLFS-ZINC34579208-1.pdb	-45.4169	-29.9432	-15.4736	0	24
VYKSPNAYTLFS-ZINC34579143-0.pdb	-48.6282	-28.5429	-20.0853	0	26
VYKSPNAYTLFS-ZINC34579142-1.pdb	-45.7017	-25.351	-20.3506	0	23.5

VYKSPNAYTLFS-ZINC34574889-1.pdb	-50.6556	-38.1462	-12.5093	0	18.5
VYKSPNAYTLFS-ZINC34574888-0.pdb	-60.9554	-46.9554	-14	0	18.5625
VYKSPNAYTLFS-ZINC34574680-1.pdb	-60.1131	-50.6453	-9.46779	0	20.1875
VYKSPNAYTLFS-ZINC34571646-0.pdb	-54.8258	-40.8309	-13.9949	0	19.3571
VYKSPNAYTLFS-ZINC34569717-0.pdb	-47.717	-38.217	-9.5	0	19.6364
VYKSPNAYTLFS-ZINC34567952-1.pdb	-56.8868	-47.1544	-9.73243	0	19.625
VYKSPNAYTLFS-ZINC34567947-1.pdb	-54.6815	-39.8485	-14.833	0	18.5294
VYKSPNAYTLFS-ZINC34567945-0.pdb	-51.0024	-42.4719	-8.5305	0	18.8125
VYKSPNAYTLFS-ZINC34567935-1.pdb	-59.8674	-40.0729	-19.7945	0	18.75
VYKSPNAYTLFS-ZINC34567932-0.pdb	-58.4604	-46.881	-11.5794	0	19.1765
VYKSPNAYTLFS-ZINC34567930-1.pdb	-47.7422	-30.0813	-18.0965	0.435573	18.4375
VYKSPNAYTLFS-ZINC34567927-1.pdb	-46.9936	-41.9991	-4.9945	0	18.625
VYKSPNAYTLFS-ZINC34567673-0.pdb	-52.0067	-43.6633	-8.34344	0	21.3077
VYKSPNAYTLFS-ZINC34567670-1.pdb	-53.5203	-43.9656	-9.55463	0	22.3846
VYKSPNAYTLFS-ZINC34567373-1.pdb	-58.0468	-34.1427	-23.9041	0	21.5833
VYKSPNAYTLFS-ZINC34567104-1.pdb	-53.4235	-32.5896	-20.834	0	20.5
VYKSPNAYTLFS-ZINC34561654-1.pdb	-51.7971	-33.0231	-18.774	0	19.8182
VYKSPNAYTLFS-ZINC34559924-0.pdb	-46.0665	-32.7299	-13.3367	0	21.2727
VYKSPNAYTLFS-ZINC34557241-1.pdb	-42.3828	-35.45	-6.93278	0	19
VYKSPNAYTLFS-ZINC34557240-0.pdb	-43.6055	-33.1055	-10.5	0	20.0833
VYKSPNAYTLFS-ZINC34557239-1.pdb	-40.8927	-33.9598	-6.93289	0	19
VYKSPNAYTLFS-ZINC34557238-0.pdb	-42.4653	-35.4653	-7	0	19.3636
VYKSPNAYTLFS-ZINC34552885-1.pdb	-45.3933	-38.4076	-6.98569	0	20.7273
VYKSPNAYTLFS-ZINC34552848-0.pdb	-57.3892	-44.5961	-12.7931	0	20.0667
VYKSPNAYTLFS-ZINC34552406-1.pdb	-54.7481	-45.2481	-9.5	0	21
VYKSPNAYTLFS-ZINC34551303-0.pdb	-54.48	-41.1148	-13.3652	0	20.1538
VYKSPNAYTLFS-ZINC34541600-0.pdb	-52.8983	-39.0032	-13.8951	0	18.6429
VYKSPNAYTLFS-ZINC34541048-1.pdb	-58.0057	-48.6838	-9.32193	0	19
VYKSPNAYTLFS-ZINC34541046-1.pdb	-60.3863	-46.3863	-14	0	18.5294
VYKSPNAYTLFS-ZINC34539694-1.pdb	-49.7451	-31.1403	-18.6049	0	22.1111
VYKSPNAYTLFS-ZINC34537793-1.pdb	-53.9113	-46.9113	-7	0	18.7333
VYKSPNAYTLFS-ZINC34537792-0.pdb	-51.6152	-42.1152	-9.5	0	18.6667
VYKSPNAYTLFS-ZINC34529374-0.pdb	-51.4012	-31.8079	-19.5933	0	22.4
VYKSPNAYTLFS-ZINC34529373-1.pdb	-50.4539	-24.5877	-25.8662	0	19
VYKSPNAYTLFS-ZINC34529002-1.pdb	-46.4895	-39.4895	-7	0	20.9091
VYKSPNAYTLFS-ZINC34529001-1.pdb	-43.3972	-37.6158	-5.78135	0	21.3636
VYKSPNAYTLFS-ZINC34527351-1.pdb	-64.4162	-41.7319	-22.6844	0	18.4444
VYKSPNAYTLFS-ZINC34516859-0.pdb	-60.1623	-48.0121	-12.1502	0	21.1429
VYKSPNAYTLFS-ZINC34514729-0.pdb	-45.7703	-38.7703	-7	0	22.9
VYKSPNAYTLFS-ZINC34514728-0.pdb	-45.9365	-38.9365	-7	0	24.2

VYKSPNAYTLFS-ZINC34514727-1.pdb	-44.509	-37.509	-7	0	23.7778
VYKSPNAYTLFS-ZINC34514726-0.pdb	-45.1049	-38.1094	-6.99556	0	23.3333
VYKSPNAYTLFS-ZINC34502526-0.pdb	-48.6625	-39.7033	-8.95924	0	18.5
VYKSPNAYTLFS-ZINC34500149-1.pdb	-63.6756	-51.129	-12.5467	0	19.3529
VYKSPNAYTLFS-ZINC34498487-0.pdb	-59.1756	-48.6374	-10.5383	0	20.4118
VYKSPNAYTLFS-ZINC34498401-1.pdb	-56.3614	-41.821	-14.5404	0	19.9231
VYKSPNAYTLFS-ZINC34498399-0.pdb	-57.851	-34.7941	-23.0569	0	18.5625
VYKSPNAYTLFS-ZINC34496883-0.pdb	-47.8991	-45.3991	-2.5	0	18.4706
VYKSPNAYTLFS-ZINC34492113-0.pdb	-52.7199	-32.3432	-20.6584	0.281662	18.5
VYKSPNAYTLFS-ZINC34492109-0.pdb	-53.8359	-40.9908	-12.8452	0	20.5714
VYKSPNAYTLFS-ZINC34492045-1.pdb	-57.4097	-48.7343	-8.67533	0	21.6667
VYKSPNAYTLFS-ZINC34490273-1.pdb	-62.4344	-50.4344	-12	0	19.1176
VYKSPNAYTLFS-ZINC34490069-1.pdb	-59.9896	-52.9896	-7	0	20.1765
VYKSPNAYTLFS-ZINC34490067-1.pdb	-59.2848	-52.6248	-6.66007	0	18.5556
VYKSPNAYTLFS-ZINC34490065-0.pdb	-63.3981	-56.3981	-7	0	21.0588
VYKSPNAYTLFS-ZINC34489224-0.pdb	-45.3769	-32.2978	-13.4567	0.377568	23.2222
VYKSPNAYTLFS-ZINC34479753-1.pdb	-52.4461	-45.4461	-7	0	20
VYKSPNAYTLFS-ZINC34477542-0.pdb	-62.7036	-52.2036	-10.5	0	20
VYKSPNAYTLFS-ZINC34477541-1.pdb	-50.2019	-34.4061	-16.0702	0.274371	20.5
VYKSPNAYTLFS-ZINC34477540-0.pdb	-47.5876	-36.367	-11.2206	0	22.7273
VYKSPNAYTLFS-ZINC34475831-0.pdb	-58.7455	-49.2455	-9.5	0	18.75
VYKSPNAYTLFS-ZINC34475830-0.pdb	-62.7724	-52.7126	-10.0598	0	20.5
VYKSPNAYTLFS-ZINC34475829-1.pdb	-59.5661	-50.0661	-9.5	0	19.1875
VYKSPNAYTLFS-ZINC34475828-1.pdb	-59.6321	-51.1405	-8.49168	0	20.3125
VYKSPNAYTLFS-ZINC34469114-1.pdb	-65.3235	-48.059	-17.2646	0	19.5556
VYKSPNAYTLFS-ZINC34464194-1.pdb	-49.7706	-36.9625	-12.8081	0	22.8
VYKSPNAYTLFS-ZINC34464193-1.pdb	-51.9414	-35.6523	-16.2891	0	25
VYKSPNAYTLFS-ZINC34464192-0.pdb	-47.6577	-33.5617	-14.0959	0	24.1
VYKSPNAYTLFS-ZINC34464191-1.pdb	-56.608	-26.3173	-30.6302	0.339485	21
VYKSPNAYTLFS-ZINC34462809-0.pdb	-46.4501	-25.2449	-21.2053	0	19.3333
VYKSPNAYTLFS-ZINC34454786-0.pdb	-44.8621	-32.3533	-12.8635	0.354718	24.1111
VYKSPNAYTLFS-ZINC34454785-1.pdb	-42.2837	-35.2837	-7	0	22.3333
VYKSPNAYTLFS-ZINC34452039-1.pdb	-57.6843	-40.7577	-16.9266	0	22.6667
VYKSPNAYTLFS-ZINC34433807-1.pdb	-51.7121	-38.2664	-13.4457	0	20.625
VYKSPNAYTLFS-ZINC34427725-1.pdb	-59.6897	-45.8133	-13.8765	0	19.7143
VYKSPNAYTLFS-ZINC34417532-1.pdb	-69.3807	-52.6936	-16.6871	0	19.5556
VYKSPNAYTLFS-ZINC34417531-1.pdb	-70.0355	-47.8721	-22.1633	0	18.9444
VYKSPNAYTLFS-ZINC34412143-1.pdb	-56.8141	-51.3822	-5.43189	0	18.8235
VYKSPNAYTLFS-ZINC34406717-1.pdb	-52.9359	-49.4359	-3.5	0	19
VYKSPNAYTLFS-ZINC34370211-0.pdb	-53.2251	-40.1289	-13.0961	0	18.75

VYKSPNAYTLFS-ZINC34370210-0.pdb	-54.5979	-40.0206	-14.5773	0	18.5
VYKSPNAYTLFS-ZINC34370209-1.pdb	-57.4226	-38.5039	-18.9187	0	19.1875
VYKSPNAYTLFS-ZINC34370208-1.pdb	-56.5261	-41.7385	-14.7876	0	18.5
VYKSPNAYTLFS-ZINC34329749-0.pdb	-56.8711	-38.5208	-18.3503	0	23.6364
VYKSPNAYTLFS-ZINC34329748-1.pdb	-65.4229	-33.093	-32.3299	0	19.6364
VYKSPNAYTLFS-ZINC34321021-0.pdb	-55.644	-45.1921	-10.4519	0	19.3333
VYKSPNAYTLFS-ZINC34321020-0.pdb	-51.5124	-41.0188	-10.4936	0	18.7333
VYKSPNAYTLFS-ZINC34321019-1.pdb	-57.0079	-46.6158	-10.3921	0	19.0625
VYKSPNAYTLFS-ZINC34321018-0.pdb	-51.8658	-41.3658	-10.5	0	18.5
VYKSPNAYTLFS-ZINC34304487-1.pdb	-51.5188	-33.0929	-18.4259	0	20.8182
VYKSPNAYTLFS-ZINC34288034-1.pdb	-42.167	-32.667	-9.5	0	25.1429
VYKSPNAYTLFS-ZINC34287807-0.pdb	-59.7291	-43.7037	-16.0253	0	18.7059
VYKSPNAYTLFS-ZINC34287806-0.pdb	-59.2658	-44.8873	-14.3786	0	19
VYKSPNAYTLFS-ZINC34287805-1.pdb	-59.7547	-37.3127	-22.6814	0.239372	18.4706
VYKSPNAYTLFS-ZINC34287804-0.pdb	-50.4024	-33.2785	-17.1239	0	18.4706
VYKSPNAYTLFS-ZINC34278228-0.pdb	-38.7919	-31.8072	-6.9848	0	25.1429
VYKSPNAYTLFS-ZINC34235675-0.pdb	-56.1387	-43.9834	-12.1552	0	21.7333
VYKSPNAYTLFS-ZINC34182444-0.pdb	-52.1436	-39.7402	-12.4034	0	18.8824
VYKSPNAYTLFS-ZINC34161425-1.pdb	-39.9679	-25.9679	-14	0	21.8571
VYKSPNAYTLFS-ZINC34160134-0.pdb	-45.693	-30.6738	-15.3941	0.374855	21.4444
VYKSPNAYTLFS-ZINC34160133-0.pdb	-45.3508	-31.6934	-14.0211	0.363644	20.5556
VYKSPNAYTLFS-ZINC34139950-1.pdb	-52.4027	-28.9016	-23.5011	0	23.2222
VYKSPNAYTLFS-ZINC34139949-1.pdb	-47.8277	-36.1509	-11.6768	0	23.5556
VYKSPNAYTLFS-ZINC34139948-1.pdb	-44.6049	-27.4852	-17.1197	0	21.4444
VYKSPNAYTLFS-ZINC34129110-1.pdb	-56.7498	-46.2532	-10.4967	0	21.2857
VYKSPNAYTLFS-ZINC34122494-0.pdb	-56.3628	-46.1463	-10.2165	0	19.5333
VYKSPNAYTLFS-ZINC34116219-0.pdb	-56.4088	-41.8374	-14.5713	0	18.5714
VYKSPNAYTLFS-ZINC34082726-0.pdb	-53.9616	-34.9616	-19	0	18.5882
VYKSPNAYTLFS-ZINC34076870-1.pdb	-64.7002	-51.7135	-12.9867	0	21.5333
VYKSPNAYTLFS-ZINC34076869-0.pdb	-57.5791	-48.0983	-9.48076	0	19.1333
VYKSPNAYTLFS-ZINC34075807-1.pdb	-56.5795	-33.2224	-23.3572	0	23
VYKSPNAYTLFS-ZINC34027303-0.pdb	-47.8933	-34.3132	-13.58	0	25.4444
VYKSPNAYTLFS-ZINC33947808-0.pdb	-59.9722	-49.9374	-10.0348	0	20.6667
VYKSPNAYTLFS-ZINC33939040-1.pdb	-60.3517	-34.6725	-25.6791	0	20.2727
VYKSPNAYTLFS-ZINC33765559-1.pdb	-63.7726	-51.7541	-12.0185	0	19.3333
VYKSPNAYTLFS-ZINC33739862-0.pdb	-46.9141	-43.4141	-3.5	0	20.5
VYKSPNAYTLFS-ZINC33693876-0.pdb	-49.0912	-42.1699	-6.92134	0	22.75
VYKSPNAYTLFS-ZINC33506585-1.pdb	-56.2692	-36.0123	-20.2568	0	20.4286
VYKSPNAYTLFS-ZINC33378647-0.pdb	-61.9392	-53.1482	-8.79091	0	20.8235
VYKSPNAYTLFS-ZINC33376898-1.pdb	-54.2697	-37.8273	-16.4423	0	22.8571

VYKSPNAYTLFS-ZINC33376445-1.pdb	-50.8112	-38.349	-12.4622	0	19.0833
VYKSPNAYTLFS-ZINC32689261-1.pdb	-50.6494	-43.6947	-6.95477	0	19.4667
VYKSPNAYTLFS-ZINC32603235-0.pdb	-51.6817	-32.7103	-18.9714	0	22.5556
VYKSPNAYTLFS-ZINC32222731-0.pdb	-49.6471	-43.4595	-6.18753	0	22.8333
VYKSPNAYTLFS-ZINC32181558-0.pdb	-35.0443	-28.0443	-7	0	25.3333
VYKSPNAYTLFS-ZINC32180809-0.pdb	-56.4687	-46.1673	-10.3014	0	20.1333
VYKSPNAYTLFS-ZINC32107279-1.pdb	-61.507	-45.5513	-15.9558	0	18.4706
VYKSPNAYTLFS-ZINC32107277-0.pdb	-64.8584	-50.8584	-14	0	19.1176
VYKSPNAYTLFS-ZINC32107226-0.pdb	-65.2239	-46.8035	-18.4204	0	21.4375
VYKSPNAYTLFS-ZINC32107224-0.pdb	-57.5354	-45.7228	-11.8126	0	19.0625
VYKSPNAYTLFS-ZINC32107130-0.pdb	-56.5202	-33.402	-23.1182	0	18.5
VYKSPNAYTLFS-ZINC31776901-1.pdb	-58.0356	-44.2085	-13.8271	0	23
VYKSPNAYTLFS-ZINC31517997-1.pdb	-61.7249	-47.7249	-14	0	18.5625
VYKSPNAYTLFS-ZINC29553629-0.pdb	-60.7952	-47.9351	-12.86	0	19
VYKSPNAYTLFS-ZINC29553625-0.pdb	-54.2368	-45.2793	-9.32326	0.365662	18.6471
VYKSPNAYTLFS-ZINC29546418-0.pdb	-39.8244	-32.8244	-7	0	22.5556
VYKSPNAYTLFS-ZINC29546415-1.pdb	-39.8699	-32.8699	-7	0	21.4444
VYKSPNAYTLFS-ZINC28826366-0.pdb	-58.9713	-56.9443	-2.02702	0	20.25
VYKSPNAYTLFS-ZINC28826350-1.pdb	-65.8386	-61.343	-4.49562	0	20.5882
VYKSPNAYTLFS-ZINC28826348-0.pdb	-58.3626	-49.2583	-9.35709	0.252819	18.7333
VYKSPNAYTLFS-ZINC28826345-1.pdb	-63.5876	-50.5955	-12.9922	0	19.2353
VYKSPNAYTLFS-ZINC28826344-0.pdb	-59.4598	-45.5682	-13.8916	0	20
VYKSPNAYTLFS-ZINC28826341-1.pdb	-56.6987	-47.8375	-8.86115	0	18.4706
VYKSPNAYTLFS-ZINC28826340-1.pdb	-55.7956	-52.4112	-3.38436	0	19.8667
VYKSPNAYTLFS-ZINC28759422-1.pdb	-53.1489	-45.4777	-7.67116	0	18.5
VYKSPNAYTLFS-ZINC28704829-1.pdb	-51.9942	-40.6594	-11.3348	0	20.5
VYKSPNAYTLFS-ZINC28100854-1.pdb	-61.2097	-34.51	-26.6997	0	18.625
VYKSPNAYTLFS-ZINC28100848-1.pdb	-63.1965	-45.6965	-17.5	0	18.5625
VYKSPNAYTLFS-ZINC28100843-1.pdb	-50.0768	-47.5768	-2.5	0	18.6875
VYKSPNAYTLFS-ZINC28100838-0.pdb	-50.1862	-37.3707	-12.8155	0	18.625
VYKSPNAYTLFS-ZINC27553251-1.pdb	-67.3421	-50.4213	-16.9208	0	19.9412
VYKSPNAYTLFS-ZINC27553248-0.pdb	-64.9099	-51.5704	-13.3395	0	19.4118
VYKSPNAYTLFS-ZINC27192423-0.pdb	-60.9165	-43.3704	-17.5461	0	18.75
VYKSPNAYTLFS-ZINC27187558-0.pdb	-57.1815	-42.0153	-15.1662	0	20
VYKSPNAYTLFS-ZINC27187553-0.pdb	-56.5711	-32.2236	-24.3475	0	18.5
VYKSPNAYTLFS-ZINC27185422-1.pdb	-60.0396	-45.0292	-15.0104	0	19.7333
VYKSPNAYTLFS-ZINC27185410-1.pdb	-55.0055	-36.8101	-18.1954	0	18.4667
VYKSPNAYTLFS-ZINC27076149-1.pdb	-61.314	-42.8193	-18.4947	0	18.5714
VYKSPNAYTLFS-ZINC26838668-1.pdb	-57.1133	-40.6868	-16.4266	0	20.25
VYKSPNAYTLFS-ZINC26838664-1.pdb	-59.997	-27.8735	-32.1235	0	19.75

VYKSPNAYTLFS-ZINC26829224-0.pdb	-59.6076	-41.8574	-17.7503	0	19.0667
VYKSPNAYTLFS-ZINC26829220-1.pdb	-59.3786	-38.4251	-20.9535	0	19.3333
VYKSPNAYTLFS-ZINC26672590-1.pdb	-65.8979	-51.0704	-14.8275	0	19.1111
VYKSPNAYTLFS-ZINC26672586-1.pdb	-67.6052	-48.2232	-19.382	0	19.5556
VYKSPNAYTLFS-ZINC26547916-0.pdb	-44.0852	-31.5297	-12.9111	0.355501	24.3333
VYKSPNAYTLFS-ZINC26547913-0.pdb	-45.3508	-31.3265	-14.4592	0.434894	19
VYKSPNAYTLFS-ZINC26547891-1.pdb	-62.4925	-42.5481	-19.9444	0	19.2857
VYKSPNAYTLFS-ZINC26547888-0.pdb	-60.1436	-35.6061	-24.5375	0	18.5333
VYKSPNAYTLFS-ZINC26546949-0.pdb	-47.7053	-40.7053	-7	0	18.4286
VYKSPNAYTLFS-ZINC26546943-1.pdb	-45.6143	-38.6143	-7	0	19
VYKSPNAYTLFS-ZINC26021496-1.pdb	-56.8729	-31.3557	-25.5171	0	19.3125
VYKSPNAYTLFS-ZINC26019861-0.pdb	-63.4081	-34.5044	-28.9038	0	19.2
VYKSPNAYTLFS-ZINC26012015-1.pdb	-60.2255	-46.3297	-13.8958	0	19.7059
VYKSPNAYTLFS-ZINC26012012-1.pdb	-61.4981	-39.6111	-21.887	0	18.5882
VYKSPNAYTLFS-ZINC26012005-1.pdb	-60.1715	-41.8705	-18.301	0	19.0588
VYKSPNAYTLFS-ZINC26011999-0.pdb	-59.9766	-38.5079	-21.4687	0	20.2353
VYKSPNAYTLFS-ZINC25721242-0.pdb	-57.4088	-41.9761	-15.4327	0	20.9286
VYKSPNAYTLFS-ZINC25627094-1.pdb	-59.6005	-48.0653	-11.5351	0	19.4706
VYKSPNAYTLFS-ZINC25626937-1.pdb	-64.2475	-47.768	-16.4795	0	18.7059
VYKSPNAYTLFS-ZINC22204334-0.pdb	-67.9142	-46.539	-21.3752	0	18.4706
VYKSPNAYTLFS-ZINC22204315-1.pdb	-65.1417	-46.3464	-18.7953	0	19.5294
VYKSPNAYTLFS-ZINC22203891-1.pdb	-62.9235	-54.2623	-8.66127	0	19.8235
VYKSPNAYTLFS-ZINC22198994-0.pdb	-56.8364	-28.0401	-28.7964	0	18.7143
VYKSPNAYTLFS-ZINC22130459-1.pdb	-66.4794	-50.8036	-15.6757	0	20.0556
VYKSPNAYTLFS-ZINC22130406-0.pdb	-58.7567	-47.6271	-11.1296	0	21.1333
VYKSPNAYTLFS-ZINC22114256-1.pdb	-51.1572	-43.0803	-8.07692	0	20.6154
VYKSPNAYTLFS-ZINC22064434-0.pdb	-59.515	-45.0644	-14.4507	0	18.8
VYKSPNAYTLFS-ZINC22064432-1.pdb	-60.3454	-46.3467	-13.9987	0	19.2
VYKSPNAYTLFS-ZINC22058744-1.pdb	-59.9075	-37.4412	-22.4662	0	18.5882
VYKSPNAYTLFS-ZINC22058740-0.pdb	-58.6471	-51.16	-7.48714	0	19.4118
VYKSPNAYTLFS-ZINC22058732-1.pdb	-60.4677	-44.6825	-15.7852	0	18.4667
VYKSPNAYTLFS-ZINC22056810-0.pdb	-52.3701	-34.7854	-17.5847	0	22.3636
VYKSPNAYTLFS-ZINC22049550-1.pdb	-54.8932	-40.3007	-14.5925	0	23.7273
VYKSPNAYTLFS-ZINC22049545-1.pdb	-52.2278	-38.1098	-14.118	0	20.4167
VYKSPNAYTLFS-ZINC22049482-1.pdb	-57.1642	-37.686	-19.4782	0	18.4706
VYKSPNAYTLFS-ZINC22049477-1.pdb	-57.2593	-37.3717	-19.8876	0	18.4706
VYKSPNAYTLFS-ZINC22049472-1.pdb	-59.5881	-45.9197	-13.6684	0	19.4706
VYKSPNAYTLFS-ZINC22049467-0.pdb	-57.5837	-39.6205	-17.9633	0	18.5882
VYKSPNAYTLFS-ZINC22049463-0.pdb	-62.0983	-44.7595	-17.3388	0	18.4706
VYKSPNAYTLFS-ZINC22049458-1.pdb	-57.8672	-44.4136	-13.4536	0	18.4706

VYKSPNAYTLFS-ZINC22049453-0.pdb	-55.8994	-40.9842	-14.9152	0	19.9412
VYKSPNAYTLFS-ZINC22049448-0.pdb	-57.842	-43.0461	-14.7959	0	20.2857
VYKSPNAYTLFS-ZINC22049444-1.pdb	-57.1273	-42.4123	-14.7149	0	20
VYKSPNAYTLFS-ZINC22049439-0.pdb	-53.5036	-41.2195	-12.6368	0.352678	18.9286
VYKSPNAYTLFS-ZINC22049416-0.pdb	-54.9318	-34.3585	-20.5733	0	21.1429
VYKSPNAYTLFS-ZINC22049412-0.pdb	-59.6842	-28.9907	-30.6936	0	19.5714
VYKSPNAYTLFS-ZINC22049403-0.pdb	-47.8125	-28.6341	-19.1784	0	25.875
VYKSPNAYTLFS-ZINC22049399-0.pdb	-53.5335	-28.1324	-25.4011	0	23.625
VYKSPNAYTLFS-ZINC22046554-0.pdb	-53.9482	-43.5389	-10.4093	0	18.4706
VYKSPNAYTLFS-ZINC22046550-1.pdb	-52.1847	-44.7578	-7.42684	0	19.3529
VYKSPNAYTLFS-ZINC22009511-0.pdb	-47.8225	-35.8296	-11.993	0	20.4545
VYKSPNAYTLFS-ZINC22002320-1.pdb	-58.8268	-46.0452	-12.7816	0	20.0667
VYKSPNAYTLFS-ZINC22002315-0.pdb	-50.4873	-38.2239	-12.2634	0	20
VYKSPNAYTLFS-ZINC22002311-1.pdb	-53.5317	-38.2134	-15.3184	0	18.9333
VYKSPNAYTLFS-ZINC22002306-0.pdb	-55.1233	-42.6433	-12.4799	0	19.8667
VYKSPNAYTLFS-ZINC21993116-1.pdb	-53.8751	-32.7426	-21.1325	0	21.8182
VYKSPNAYTLFS-ZINC21990362-1.pdb	-50.9287	-45.5609	-5.36783	0	19.0667
VYKSPNAYTLFS-ZINC21990359-1.pdb	-59.0912	-44.9025	-14.1887	0	19.5333
VYKSPNAYTLFS-ZINC21990356-0.pdb	-54.565	-52.065	-2.5	0	18.8235
VYKSPNAYTLFS-ZINC21990354-1.pdb	-58.4766	-52.4766	-6	0	18.4706
VYKSPNAYTLFS-ZINC21989403-1.pdb	-49.5877	-42.5877	-7	0	18.4667
VYKSPNAYTLFS-ZINC21989401-1.pdb	-49.7468	-42.7468	-7	0	18.4667
VYKSPNAYTLFS-ZINC21989398-1.pdb	-49.1666	-49.1666	0	0	18.6667
VYKSPNAYTLFS-ZINC21989396-1.pdb	-48.8228	-39.9819	-8.84095	0	18.5333
VYKSPNAYTLFS-ZINC21985602-0.pdb	-61.1329	-30.3022	-30.8307	0	19.0909
VYKSPNAYTLFS-ZINC21985597-1.pdb	-55.7793	-46.2793	-9.5	0	18.9286
VYKSPNAYTLFS-ZINC21985368-1.pdb	-46.8542	-45.7367	-1.11746	0	20.5
VYKSPNAYTLFS-ZINC21985366-1.pdb	-49.3807	-48.3366	-1.04408	0	21.2857
VYKSPNAYTLFS-ZINC20426942-1.pdb	-55.8461	-44.0007	-11.8455	0	19.7857
VYKSPNAYTLFS-ZINC20232286-0.pdb	-58.4186	-28.3646	-30.054	0	18.4615
VYKSPNAYTLFS-ZINC20232281-1.pdb	-64.0325	-34.3518	-29.6807	0	19.3077
VYKSPNAYTLFS-ZINC19907272-1.pdb	-51.2166	-44.2166	-7	0	19.3333
VYKSPNAYTLFS-ZINC19893242-0.pdb	-59.9221	-54.2077	-5.71446	0	19.2941
VYKSPNAYTLFS-ZINC19891524-1.pdb	-59.1218	-52.2758	-6.846	0	19.8824
VYKSPNAYTLFS-ZINC19872842-0.pdb	-48.0035	-39.5006	-8.5029	0	20.1667
VYKSPNAYTLFS-ZINC19872840-1.pdb	-49.6994	-43.5221	-6.17738	0	22.8333
VYKSPNAYTLFS-ZINC19872837-1.pdb	-57.0293	-51.9532	-5.07608	0	18.5
VYKSPNAYTLFS-ZINC17724845-0.pdb	-58.8854	-36.4528	-22.4326	0	18.7059
VYKSPNAYTLFS-ZINC17421447-1.pdb	-49.6549	-47.1549	-2.5	0	19
VYKSPNAYTLFS-ZINC17378346-0.pdb	-45.988	-28.8378	-17.1502	0	21.8

VYKSPNAYTLFS-ZINC17314801-1.pdb	-56.9317	-45.9317	-11	0	20.3571
VYKSPNAYTLFS-ZINC17300590-0.pdb	-57.9089	-41.2597	-16.6492	0	20.0769
VYKSPNAYTLFS-ZINC17288490-1.pdb	-54.8833	-45.3833	-9.5	0	21.7692
VYKSPNAYTLFS-ZINC17197391-0.pdb	-60.3834	-31.7644	-28.9382	0.319212	19
VYKSPNAYTLFS-ZINC17176129-0.pdb	-48.4032	-41.6407	-6.76253	0	19.5714
VYKSPNAYTLFS-ZINC17175346-0.pdb	-49.5346	-33.0346	-16.5	0	20
VYKSPNAYTLFS-ZINC17175261-1.pdb	-49.5511	-40.1477	-9.40337	0	24.1
VYKSPNAYTLFS-ZINC17173515-1.pdb	-54.6811	-34.4521	-20.5052	0.276203	23.8
VYKSPNAYTLFS-ZINC17164548-0.pdb	-57.6872	-51.9364	-5.75077	0	20.1667
VYKSPNAYTLFS-ZINC17161511-0.pdb	-49.6984	-37.0476	-12.6508	0	19.6923
VYKSPNAYTLFS-ZINC17108604-1.pdb	-57.3681	-39.5468	-17.8213	0	22.2727
VYKSPNAYTLFS-ZINC17064522-1.pdb	-57.9748	-50.9748	-7	0	19.6875
VYKSPNAYTLFS-ZINC17063817-0.pdb	-55.7344	-43.8036	-11.9309	0	20.0714
VYKSPNAYTLFS-ZINC17063815-0.pdb	-60.1013	-43.2365	-16.8648	0	21.0714
VYKSPNAYTLFS-ZINC17047258-1.pdb	-58.3489	-43.3784	-14.9705	0	19.6429
VYKSPNAYTLFS-ZINC17001182-1.pdb	-55.4565	-40.9074	-14.5491	0	21.8462
VYKSPNAYTLFS-ZINC16997856-1.pdb	-49.1275	-32.4968	-16.9452	0.314591	20.4545
VYKSPNAYTLFS-ZINC16982373-0.pdb	-65.1802	-56.8691	-8.31107	0	21.3889
VYKSPNAYTLFS-ZINC16968896-0.pdb	-55.1697	-50.0923	-5.07744	0	20.7857
VYKSPNAYTLFS-ZINC16955443-1.pdb	-55.8506	-41.8508	-13.9998	0	20.8462
VYKSPNAYTLFS-ZINC16954002-1.pdb	-43.4843	-39.9886	-3.49564	0	19.5385
VYKSPNAYTLFS-ZINC16952697-0.pdb	-46.2516	-42.7516	-3.5	0	18.4706
VYKSPNAYTLFS-ZINC16952695-1.pdb	-49.7854	-46.3098	-3.47561	0	19.5333
VYKSPNAYTLFS-ZINC16941446-0.pdb	-53.9915	-40.9138	-13.0777	0	19.3125
VYKSPNAYTLFS-ZINC16941442-1.pdb	-51.6702	-38.1111	-13.5591	0	21.3077
VYKSPNAYTLFS-ZINC16940978-1.pdb	-44.722	-25.1469	-19.5751	0	20.625
VYKSPNAYTLFS-ZINC16940076-1.pdb	-56.3418	-41.3814	-14.9604	0	20.6667
VYKSPNAYTLFS-ZINC16923303-0.pdb	-51.0758	-48.5812	-2.49463	0	18.8125
VYKSPNAYTLFS-ZINC16696935-1.pdb	-56.3467	-46.8853	-9.46141	0	20.0588
VYKSPNAYTLFS-ZINC16696932-0.pdb	-55.2488	-44.152	-11.0968	0	20.75
VYKSPNAYTLFS-ZINC16696930-0.pdb	-62.1698	-45.3272	-16.8426	0	19.5882
VYKSPNAYTLFS-ZINC16123805-1.pdb	-51.2624	-38.6561	-12.6064	0	20.9167
VYKSPNAYTLFS-ZINC15937364-0.pdb	-56.0892	-46.5892	-9.5	0	18.5
VYKSPNAYTLFS-ZINC15937228-1.pdb	-52.7566	-45.8227	-6.93396	0	18.7143
VYKSPNAYTLFS-ZINC15937227-0.pdb	-52.3191	-36.9014	-15.6965	0.27889	18.5
VYKSPNAYTLFS-ZINC15935031-1.pdb	-54.155	-42.8319	-11.3231	0	18.5625
VYKSPNAYTLFS-ZINC15922260-1.pdb	-55.2799	-45.8123	-9.46764	0	18.5294
VYKSPNAYTLFS-ZINC15894679-1.pdb	-61.4056	-43.9606	-17.445	0	20.8
VYKSPNAYTLFS-ZINC15785332-1.pdb	-52.3768	-45.3768	-7	0	18.5333
VYKSPNAYTLFS-ZINC15785118-0.pdb	-52.4424	-37.7134	-14.7289	0	21.5455

VYKSPNAYTLFS-ZINC15784788-1.pdb	-48.6639	-36.252	-12.4119	0	19.6923
VYKSPNAYTLFS-ZINC15784786-0.pdb	-44.626	-33.453	-11.173	0	22.5
VYKSPNAYTLFS-ZINC15784784-1.pdb	-46.8674	-40.776	-6.09139	0	23.2727
VYKSPNAYTLFS-ZINC15784782-1.pdb	-46.434	-40.5153	-5.91874	0	23.3636
VYKSPNAYTLFS-ZINC15784780-1.pdb	-47.2555	-41.2121	-6.04336	0	23.3636
VYKSPNAYTLFS-ZINC15784546-0.pdb	-58.2402	-52.5056	-5.73467	0	20.4375
VYKSPNAYTLFS-ZINC15784544-0.pdb	-55.6665	-46.2791	-9.38742	0	21.0714
VYKSPNAYTLFS-ZINC15784358-0.pdb	-51.9988	-34.3997	-17.5991	0	20.8
VYKSPNAYTLFS-ZINC15784356-1.pdb	-50.8377	-47.3377	-3.5	0	20.9231
VYKSPNAYTLFS-ZINC15784354-1.pdb	-49.6606	-43.421	-6.2396	0	22.6667
VYKSPNAYTLFS-ZINC15783176-1.pdb	-55.645	-48.7516	-6.89332	0	20.625
VYKSPNAYTLFS-ZINC15782452-0.pdb	-53.6382	-34.0165	-19.6217	0	22.4545
VYKSPNAYTLFS-ZINC15781858-1.pdb	-64.1145	-50.1472	-13.9673	0	22.1429
VYKSPNAYTLFS-ZINC15781121-0.pdb	-55.4475	-41.9497	-13.4978	0	20.4615
VYKSPNAYTLFS-ZINC15780430-1.pdb	-66.5006	-40.3675	-26.1331	0	18.6
VYKSPNAYTLFS-ZINC15780220-0.pdb	-64.5033	-53.3748	-11.1285	0	19.1765
VYKSPNAYTLFS-ZINC15780218-0.pdb	-56.7917	-47.4859	-9.3058	0	20.5
VYKSPNAYTLFS-ZINC15779471-1.pdb	-56.3374	-42.3478	-13.9896	0	20.4167
VYKSPNAYTLFS-ZINC15778754-1.pdb	-56.5915	-46.0934	-10.4981	0	19.6
VYKSPNAYTLFS-ZINC15778697-0.pdb	-63.4126	-45.515	-18.1577	0.260111	19.4
VYKSPNAYTLFS-ZINC15443135-0.pdb	-51.5021	-37.5021	-14	0	20.9091
VYKSPNAYTLFS-ZINC15443133-1.pdb	-48.8158	-34.8668	-13.949	0	21.3636
VYKSPNAYTLFS-ZINC15443126-0.pdb	-62.1473	-32.5383	-29.609	0	20.0909
VYKSPNAYTLFS-ZINC15440710-1.pdb	-47.9766	-40.9766	-7	0	19.6923
VYKSPNAYTLFS-ZINC15440708-1.pdb	-45.0148	-38.0352	-6.97963	0	19.0769
VYKSPNAYTLFS-ZINC15022795-0.pdb	-53.7064	-40.5723	-13.134	0	23.0769
VYKSPNAYTLFS-ZINC15022783-0.pdb	-54.0612	-43.571	-10.4902	0	18.6

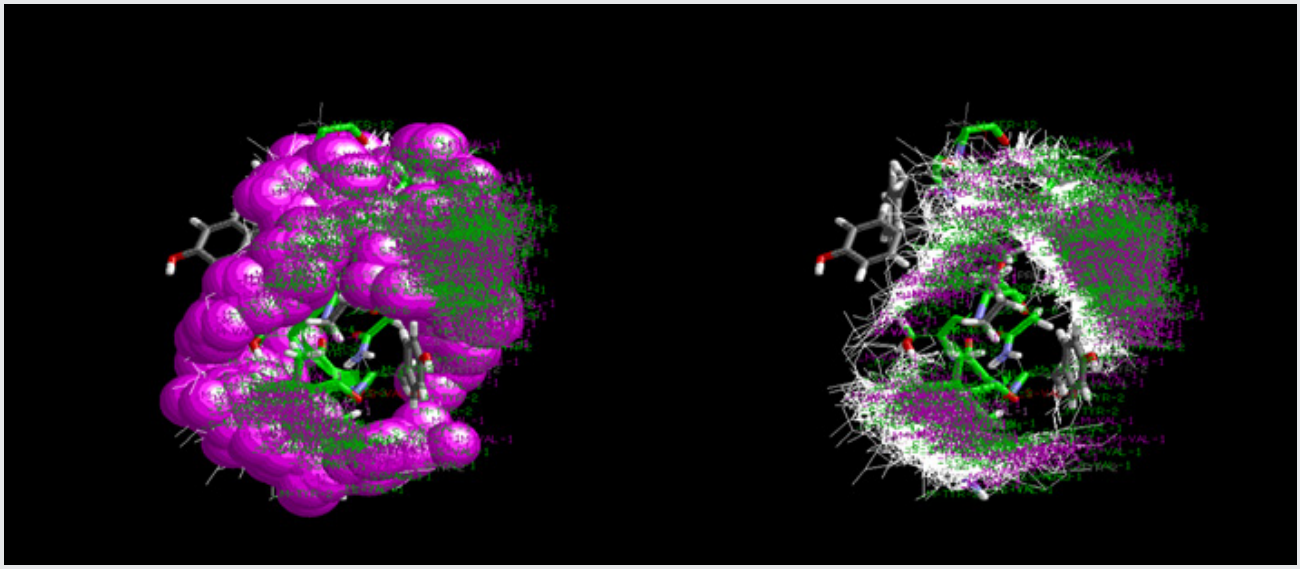


Figure 28

Table 11 represent the MM-PBSA limit when we try to compare the binding affinities of different complexes, mainly when dealing with docking since the conformational sampling is poor with respect to performing MM-PBSA from molecular dynamics simulations (Spiliotopoulos et al., 2012). Nevertheless, in case of AIRE-PHD1 mutants, the MM-PBSA and dMM-PBSA uncertainty over the NMR structure bundle (20 structures) reduces, leading to an average standard error of 34 and 7 kJ/mol, respectively, indicating that dMM-PBSA is fairly reliable in predicting protein-peptide interface alanine mutations. Moreover, the CAS dMM-PBSA of AIRE-PHD1 error range values are fairly in agreement with the CAS MM-PBSA error range (ca. 4 kJ/mol) shown in Spiliotopoulos et al., in which the sampling was carried out by molecular dynamics simulations (Spiliotopoulos et al., 2012). This result indicates that dMM-PBSA carried out on a small number of structures (e.g., 20) behaves similarly to MM-PBSA from molecular dynamics, in which the sampling is more extended. In fact, the modulation of the polar terms in MM-PBSA is probably taking into account the possible local rearrangement similarly to what is done via a higher internal dielectric, indicating that dMM-PBSA represents a simple and promising approach in evaluating alanine mutations from single structure, either from docking or from X-Ray or from NMR. We believe that in parallel with the recently developed optimization of the HADDOCK score for PPIs inhibitors (Kastritis et al., 2014), the combination of HADDOCK score and modified MM-PBSA binding free energy might lay the groundwork for novel approaches to study in silico PPIs inhibitors in a quick and automatic fashion. The advantage of using MM-PBSA as scoring function is threefold. First, MM-PBSA is versatile. It provides estimates of the equilibrium averages over the solvent degrees of freedom, permitting the post-processing of solute representative snapshots from docking poses. Since MM-PBSA estimates binding free energy, it represents a valuable alternative since it is in principle transferable between different docking runs and can be used to score both intra-ligand and inter-ligand poses, saving individual validation for each system under study. This makes MM-PBSA more suitable for novel problems with limited experimental data as we demonstrated in the case study of NPHP1-SH3. Second, MM-PBSA can be used to quantify the thermodynamical strength of the putative poses. In particular, our dMM-PBSA is a reliable scoring function in the protein-peptide field showing good a correlation with experimental data. Therefore, the advantage to use dMM-PBSA with respect to MM-PBSA relies on the possibility to modulate the polar terms without rerunning the Poisson-Boltzmann calculations at different internal dielectric values. Finally, MM-PBSA could better allow disclosing atomistic details of protein-peptide binding, supporting the rational design of bioactive compounds. In fact, post-processing task such as CAS approach has been very recently applied to MD simulations for successfully evaluating the importance of key residues in the protein-peptide binding complex (Spiliotopoulos et al., 2012). Therefore, MM-PBSA can be used to calculate the ΔΔG, defined as ΔGwt - ΔGmutALA, on the protein-peptide best docked poses in order to identify residues for which mutation to alanine strongly attenuates binding. The latter behavior occurred in our short CAS study of AIRE-PHD1 mutants Figure 3, by which an acceptable correlation of R = 0.86/0.88 between the experimental and the calculated binding energies of the mutants demonstrated the possibility to use MM-PBSA as a promising tool at low computational cost to evaluate the hot-spots in the PPIs field with respect to the HADDOCK score. Finally, structural prediction of protein-peptide complexes remains challenging due to two major obstacles: peptides are highly flexible and they often interact weakly with their substrate, underlining their importance in signal transduction or regulation which often relies on transient processes. This leaves flexible docking as one of the few amenable computational techniques to model these complexes (Verkhivker et al., 2000; Hetényi and van der Spoel, 2002; Niv and Weinstein, 2005; Raveh et al., 2010, 2011; Trellet et al., 2013). In our study, we are considering the two interacting partners, protein and peptide, already in the bound conformation, which represents a strong assumption in term of binding mechanism and also the best scenario for a docking calculation, especially in the PPIs field. This undoubtedly increases the success rate in the native pose determination and it reduces the error in the binding affinity calculations since in our MM-PBSA approach we are neglecting most of the solute entropic contribution (i.e., under this assumption the estimate of ΔS = SComplex - (Sprotein - Speptide) can be poor). Recent work have highlighted the efforts to improve HADDOCK protocol in the field of protein-peptide (Trellet et al., 2013) but still predicting large conformational changes remains a challenge as indicated by several failure to accurately predict cases where the protein undergoes large conformational changes upon binding (Trellet et al., 2013). In this case, even the more accurate and reliable scoring function and binding free energy methods will struggle in discriminating the correct binding mode due to both hard and soft docking failure (Verkhivker et al., 2000). In conclusion, in contrast to other scoring functions and approximate binding free energy calculation methods such as the linear interaction energy (LIE) method, MM-PBSA contains less empirical parameters and, thus, it is more likely to be useful in determining the relative free energies of binding of quite different compounds and systems for which there is more limited experimental data, although a protein structure of the target is, of course, required. DG based solvation models have had a considerable success in solvation analysis 10,11,13,70. Particularly, our DG based non-polar solvation model was shown to offer some of the most accurate solvation energy predictions of various non-polar molecules.16 However, our DG based full solvation model is subject to numerical instability in solving the GLB equation, due to its coupling with the GPB equation. To stabilize the coupled GLB and GPB equations, a strong constraint on the van der Waals interaction was applied in our earlier work, 10,11,13 which hinders the parameter optimization of our DG based solvation model. In the present work, we resolve this problem by introducing new parameter optimization algorithms, namely, perturbation method and convex optimization, for the DG based solvation model. New stability conditions are explicitly imposed to the parameter selection, which guarantees the stability and robustness of solving the GLB equation and leads to constrained optimization of our DG based solvation model. The new optimization algorithms are intensively validated by using a large number of test molecules, including the SAMPL0 test set, 53 alkane, alkene, ether, alcohol, and phenol types of solutes. Regression results based on our new algorithms are consistent extremely well with experimental data. Additionally, a fivefold cross validation technique is employed to explore the ability of our DG based solvation models for the blind prediction of the solvation free energies for a variety of solute molecules. It is found that the same level of errors is found in the training and validation sets, which confirms our model's predictive power in solvation free energy analysis. The present DG based full solvation model provides a unified framework for analyzing both polar and nonpolar molecules. In our future

work, we will develop machine learning approaches for the robust clustering of solute molecules of interest into appropriate categories so as to better predict their solvation free energies. The conformational space and the dynamics associated with macromolecules and macromolecular complexes, as well as their structural stability, may be explored and asserted through molecular dynamics and Monte Carlo simulations. Although full knowledge of the algorithms involved is not absolutely required, to obtain meaningful simulations, it is impressible to understand the conditions, approximations, and hypotheses on which they are based, as well as their limitations and potentialities. Although most simulations are currently limited to individual structures or small organisms and complexes (a few millions atoms), it is conceivable that in the near future large viruses, such as the influenza, may be entirely simulated in silico over long periods of time. That would open the door to an essentially in silico design and lead to a better assessment of toxicity, specificity, and efficiency. Such a phenomenon has occurred before in aerodynamics in which computational fluid dynamics has essentially replaced wind tunnel experiments in the design, of large aeroplane fuselages and wings. In this paper, we present and compare different parallel implementations of a full PBE solver based on a finite-difference scheme. As for the algorithm itself, we follow the approach of the DelPhi PBE solver, which exploits the checkerboard structure of the finite difference discretization of the Laplace differential operator and adopts a successive overrelaxation scheme to converge to the solution. We parallelize the algorithm using OpenMP, MPI, and CUDA to exploit multicore CPUs, clusters of multicore CPUs, and GPUs. A MPI-CUDA implementation was used to exploit clusters of GPUs. The MPI implementation achieved good speedup values, up to 30 times the serial code using 50 cores. When compared with OpenMP or OpenMP and MPI used together, MPI showed better performance. Many different points can be investigated as future directions, as the use of nonblocking communication and the features present in CUDA 5 (and onward). The algorithm can also be improved looking for cache optimization, so that the memory access on the Boltzmann stencil is done in a more intelligent way. This paper gives an overview of the mathematical methods currently used in quantitative structure-activity/property relationship (QASR/QSPR) studies.Recently, the mathematical methods used as regression tools in QSAR/QSPR analysis have been developing quickly. Thus, not only are the previous methods, such as Multiple Linear Regression (MLR), Partial Least Squares (PLS), Neural Networks (NN), Support Vector Machine (SVM), being upgraded by improving the kernel algorithms or by combining them with other methods, but also some new methods, including Gene Expression Programming (GEP), Project Pursuit Regression (PPR) and Local Lazy Regression (LLR), are being mentioned in the current reported QSAR/QSPR studies. In light of this, this paper is divided as follows: first, the improved methods based on the earlier ones are described, then the new methods are presented. Finally, all these methods are jointly commented and their advantages and disadvantages discussed to show their application potential in future QASR/QSPR studies. We propose a new probability algorithm named EPP for identifying motifs in DNA datasets. EPP presents a new entropy-based position projection to divide original dataset and remove a large amount of redundant information. Experimental results show that EPP is able to efficiently and effectively identify motifs in DNA sequences and ChIPseq datasets. However, the functions of some motifs are still unknown; the analysis of motifs in these complex transcriptional regions is needed. In addition, with the increase of data size, designing the parallel algorithm to handle big data is also a key issue for the future study. The conformational space and the dynamics associated with macromolecules and macromolecular complexes, as well as their structural stability, may be explored and asserted through molecular dynamics and Monte Carlo simulations. Although full knowledge of the algorithms involved is not absolutely required, to obtain meaningful simulations, it is impressible to understand the conditions, approximations, and hypotheses on which they are based, as well as their limitations and potentialities. Although most simulations are currently limited to individual structures or small organisms and complexes (a few millions atoms), it is conceivable that in the near future large viruses, such as the influenza, may be entirely simulated in silico over long periods of time. That would open the door to an essentially in silico design and lead to a better assessment of toxicity, specificity, and efficiency. Such a phenomenon has occurred before in aerodynamics in which computational fluid dynamics has essentially replaced wind tunnel experiments in the design of large aeroplane fuselages and wings. The modern in silico (computational) chemical analysis respecting the bio- activity and availability of analogues substances, potentially beneficial or detrimental for specific interaction in organs and organisms, faces with a paradoxical dichotomy: if searching for the best correlation useful for prediction of specific molecular bio- or eco- activity QSAR models involving un-interpretable many latent variables may be produced, while always remaining the question of correlation factor indeterminacy (i.e. the assumed descriptors can be at any time replaced with other producing at least the same correlation performances); instead, when restricting the analysis to search for molecular design and mechanisms throughout performing SARs by means of special structural indicators for a given class of relevant molecules, arises the price of limiting the use of generated models for further prediction. The present communication is mainly devoted in developing the second (Q)SAR facet by extending the recent introduced notion of spectral-path-linking-endpoints and the associate least action principle to spectral path quantification, in terms of the best fitted molecules, along the contained computed models, by means of the introduced q(uantum)-SAR factor within the generally called Quantum-SAR (QuaSAR) methodology. As an application, for representative flavonoids' inhibiting activities on breast cancer resistant protein there was clearly shown that the newly introduced q-SAR factor offers relevant analytical characterization of previously conceptually introduced spectral path hierarchy; moreover, the present QuaSAR may allow interpretation inter-conversion of concerned molecules' towards receptor binding since belonging to the same class of analogs, while they certainly undertaking such transformation during their interaction with macromolecules, proteins and enzymes present on cellular walls or with *in vivo* environment. Basically, the QuaSAR stands as the first step in assessing the quantum mechanically equivalent of wave function to the sample of molecules interacting with a specific organism site; it will eventually lead with the hyper-wave function with the help of which the associate hyper-density probability of binding (metabolization) is to be computed; the last information may provide the density probability map of the ligand-receptor interaction abstracted from the structural Spectral-Qua-SAR correlations; with this tool the molecular design of new chemical structures may be appropriately undertaken. However, the actual QuaSAR scheme and quantum factor carry the main features of quantum dynamical systems and may stimulate future computational and conceptual developments in

molecular design for structurally controlled activity. Further generalization of the present QuaSAR method to modeling all potential inter-conversions of employed molecules involved in correlation as well as for establishing their quantum metabolization complete map (through, for instance, hydrophobic, electrostatic and steric barrier tunneling) is actually in progress and will be reported in subsequent communications. From this study some useful conclusions are summarized as follows.

i. Ten of the twenty natural amino acids are involved in the three types of strong interactions (salt-bridge, cation-π and amide bridge), which are much stronger than typical hydrogen bonds and often play important roles in protein-protein, protein-peptide, protein-ligand, and protein-DNA interactions.

ii. The salt-bridge interactions between acidic (Glu- and Asp-) and basic (Arg+, Lys+ and His+) amino acids are the strongest residue-residue interactions. However, salt-bridges may be weakened by solvation effects and may be broken by acidic conditions.

iii. The cation-π interactions between protonated amino acids (Arg+, Lys+, and His+) and aromatic amino acids (Phe, Tyr, Trp and His) are 2.5 to 5-fold stronger than typical hydrogen bonds and are less affected by solvent and pH than are salt bridge interactions. The cation-π interactions could occur at the hydrophobic-hydrophilic interface.

iv. Amide bridge interactions are special amino acid interactions that only occur between two amide amino acids (Asn and Gln), and these interactions are three times stronger thanA Copula Based Approach for Design of Multivariate Random Forests for Drug Sensitivity Prediction. Conclusions our approach enables the construction of frames with flexible evolution of time-frequency resolution over time or frequency. The resulting frames are well suited for applications as they can be implemented using fast algorithms, at a computational cost close to standard Gabor frames. Exploiting evolution of resolution over time, the proposed approach can be of particular interest for applications where the frequency characteristics of the signal are known to evolve significantly with time. Order analysis [114], in which the signal analyzed is produced by a rotating machine having changing rotating speed, is an example of such an application. Exploiting evolution of resolution over frequency, the presented approach is valuable for applications requiring the use of a tailored non uniform filter bank. In particular, it can be used to build filter banks following some perceptive frequency scale; see e.g. (68). In the present contribution, we described in detail an invertible constant-Q transform. One difficulty when using our approach is to adapt the time-frequency resolution to the evolution of the signal characteristics. If prior knowledge is available, this can be done by hand. An automatic adaptation algorithm based on onset detection was described in Section 4.1.2. A different approach will involve the investigation of sparsity criteria as proposed in [82]. Finally, future work will lead to adaptability in both time and frequency leading to quilted frames as introduced in [119]. In this paper, a new method namely, the SEG iterative method was formulated for solving the elliptic boundary value problems. The presented results show that the proposed method is capable of approximating the solution very well in terms of accuracy and execution time. It can be seen that the computation cost is reduced substantially compared to those obtained by the cubic spline block Gauss-Seidel iterative method [169], especially when the grid size increases. Furthermore, in terms of accuracy, the proposed method is superior to the original central difference explicit group iterative method [164]. In conclusion, the proposed method is a viable alternative approximation tool for solving the elliptic partial differential equations. We evaluate four different ways of computing jump coefficients between voxels for stochastic simulation of diffusion in a Cartesian mesh; as a distinctive feature we include jumps along the diagonals in the mesh. The method based on the first exit time (FET) is equivalent to a finite difference method (FDM) of second order accuracy if the parameters are chosen appropriately. The most accurate and efficient method compared to the solution of the diffusion equation on a square is to use the coefficients of the finite volume method (FVM) or equivalently a FDM with jumps in the coordinate directions without including the diagonals. If we count the number of jumps to obtain a certain relative error, the differences between the methods are minor. In this paper, we focus on two algorithms solved a class of large-scale nonlinear equations. At the first step, a CG algorithm, called a sub-algorithm, was used as the initial points of the main algorithm. Then a quasi-Newton algorithm with the initial points done by a CG sub-algorithm was defined as the main algorithm. In order to avoid computing the Jacobian matrix, a nonmonotone line search technique was used in the algorithms. The convergence results are established and numerical results are reported.According to the numerical performance, it is clear that the CG technique is very effective for large-scale nonlinear equations. This observation inspires us to design the CG methods to directly solve nonlinear equations in the future. Signal analysis with classical Gabor frames leads to a fixed time-frequency resolution over the whole time-frequency plane. To overcome the limitations imposed by this rigidity, we propose an extension of Gabor theory that leads to the construction of frames with time-frequency resolution changing over time or frequency. We describe the construction of the resulting nonstationary Gabor frames and give the explicit formula for the canonical dual frame for a particular case, the painless case. We show that wavelet transforms, constant-Q transforms and more general filter banks may be modeled in the framework of nonstationary Gabor frames. Further, we present the results in the finite-dimensional case, which provides a method for implementing the above-mentioned transforms with perfect reconstruction. Finally, we elaborate on two applications of nonstationary Gabor frames in audio signal processing, namely a method for automatic adaptation to transients and an algorithm for an invertible constant-Q transform. In this short review we have raised the question of replacing commonly used ODE models for protein-protein interactions by reaction-diffusion models which are well studied [324,325,327] which not only contain reaction terms coming from the protein interactions but also describe the spatial distribution of species involved in the reactions over the cellular compartments. We have proposed the simpler Leloup-Goldbeter circadian model for (β2GPI) antibodies, Anti-phospholipid antibodies (aPL) in Neurospora that contains two equations for the nuclear and cytoplasmic (β2GPI) antibodies, Anti-phospholipid antibodies (aPL) protein and its mRNA, and a more complicated model for apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b response to the occurrence

of DSBs in a single cell. These models, and more generally reaction-diffusion PDE models for intracellular protein dynamics, are likely to be more realistic than ODE models. As an example of such possibly more realistic models, let us consider the slightly more complicated PER model in Drosophila of Leloup and Goldbeter [338,339] that relies on the same principles as the (β2GPI) antibodies, Anti-phospholipid antibodies (aPL) model. Indeed, two additional phosphorylations are considered, and this first PER model (more sophisticated models of PER have been published by Leloup and Goldbeter) is proposed to be amenable to describe two Drosophila mutants for PER, with shorter or longer period of oscillations, by lower and higher values, respectively, of the maximum degradation rate vd of cytoplasmic diphosphorylated PER. But is this sole change of constants the best reason, in a more physiological model, for the change in periods? Could it be related to transcription of genes, to species diffusion, to translation into protein products, to cytosolic/nucleic degradation, to nucleocytoplasmic transport, to the extension of the "dead zone" around the nucleus where translation occurs? The same questions can be posed when one considers apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b and drugs amenable to electively modify specific parts of its intracellular dynamics. This can be rendered optimally in PDE models, that are much more physiological than (rougher) ODE models, since they are naturally able to take intracellular spatial features into account, including possible space heterogeneities in the intracellular medium (not considered in the examples, apart from the above mentioned dead zone of non-translation in the cytoplasm). The oscillatory dynamics of apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b, "the guardian of the genome", has long been evidenced by biological recordings in experimental conditions representing cell stress due to, e.g. radiotoxic insult [158,169]. In this sense, the biological question of identifying intracellular spatio-temporal dynamics is certainly crucial. Furthermore, since apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b disruptions are found in more than 50% of solid tumours, with various modifications of its stress-induced dynamics [65,66,132], it is also a crucial point in cancer therapy modelling to understand how apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b dynamics may be affected in disease, and to this aim, to have an accurate representation of how this dynamics works in physiological conditions, which is the purpose of designing apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b dynamic models. We have shown how easy it is in principle to start from an ODE compartmental model, add diffusion terms - which in their simplest version are mere laplacians - and slightly modify the representation of exchanges between the compartments to adapt them to the new setting, to obtain a ready-made reaction-diffusion PDE model. However, dependence on spatial patterns, such as oscillations in concentration, on the diffusivities and the (nuclear membrane) permeabilities has to be better studied, with precise identification of the underlying biological parameters, to understand their roles properly. To obtain more realistic PDEs one can consider effects of different viscosities of buffers on different diffusivities of molecules or adjunction of advection terms when knowledge of aided transport makes it relevant (e.g., transport through microtubules). To support these ideas, different mobility of molecules in the nucleus and the cytoplasm caused by limited diffusions has led to biphasic caspase activation kinetics, see [140] and citations therein. In some circumstances it might be convenient to include specific spatial structures, such as position and density of ribosomes and thus to model gene expression more specifically [141,142]. In addition, intracellular signalling in our rather simplified modelling setting is restricted to cells of physiology and morphology where a molecular network in question is, at least, partially understood. What does the apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b signal transduction in response to DNA damage in a nerve cell or in a polynucleic muscle cell look like? What is the role of diffusion in a cell where a signal is spread over long distances and in very small cells? These and other questions should be further addressed and spatial PDE models might fruitfully be used for this purpose. Note also that the oscillatory patterns are self-organised not only due to events occurring in the nucleus or in the cytoplasm but that they are also tightly connected to boundary conditions on the membranes. Thus physiological delays maintained due to the semipermeability of the nuclear membrane are more typical for PDEs than for ODEs (if one does not want to deal with artificial delays represented by delay differential equations). We examined the Kedem-Katchalsky as representing passive transport mechanisms with the difference of concentrations at both sides of the membrane as the driving force for exchanges; however, bigger species are rather transported actively, which should be taken into account in more sophisticated models of nucleocytoplasmic transport [123]. On a more technical note, we have also briefly introduced the semi-implicit Rothe method that can be used for numerical solution of reaction-diffusion systems as an alternative to other used methods; they can be used also in 3D simulations, [82], where other methods are in general more demanding in time and memory. We hope that the simple presentation of this alternate solution to classical ODEs will be of some help to biologists and modellers who want to describe intracellular spatio-temporal dynamics of proteins in a faithful, yet more demanding in terms of parameter estimation, way. We would as well like to mention that such reaction-diffusion PDE models are also amenable to describe spatio-temporal dynamics at the level of cell populations and that, by introducing intercellular signalling, it is in principle possible to connect the two observation levels. This perspective still remains a challenge to mathematicians and modellers in biology.In this paper, a new method namely, the SEG iterative method was formulated for solving the elliptic boundary value on a chemoproteomic-aided Molecular designed, ovarian stem cells (OSCs) motif VYKSPNAYTLFS derived post-trancripts domains to Germline low energetics drug design problems. The presented results show that the proposed method is capable of approximating the solution very well in terms of accuracy and execution time. It can be seen that the computation cost is reduced substantially compared to those obtained by the cubic spline block Gauss-Seidel iterative method [249], especially when the grid size increases. Furthermore, in terms of accuracy, the proposed method is superior to the original central difference explicit group iterative method [318]. In conclusion, the proposed method is a viable alternative approximation tool for solving the elliptic partial differential equations. In this paper, we focus on the current mathematical methods used as regression tools in recent QSAR/QSPR in silico Reaction-diffusion Chemoproteomic-aided Molecular designed, ovarian stem cells (OSCs) motif

VYKSPNAYTLFS derived and a quasi-Newton algorithm Scan Predicted antigenic peptide-like polypharmacophoric ligand (CAMSPCPL) targeted on the conserved (β2GPI) antibodies, Anti-phospholipid antibodies (aPL) or from (PAPP-A; gi:38045915) and APOH apolipoprotein H post-trancripts domains to Germline low energetic drug discovery studies. Mathematical regression methods are so important for the QSAR/QSPR modeling that the choice of the regression method, most of the time, will determine if the resulted model will be successful or not. Fortunately, more and more new methods and algorithms have been applied to the studies of QSAR/QSPR, including linear and nonlinear, statistics and machine learning. At the same time, the existing methods have been improved. However, it is still a challenge for the researchers to choose suitable methods for modeling their systems. This paper may give some help on the knowledge of these methods, but more practical applications are needed so as to get a thorough understanding and then perform a better mathematical-aided drug design application.The modified Poisson-Nernst-Planck equations account for steric effects and lead to physically realistic results. In this paper we develop a finite element method for the MPNPE, with a focus on computing steady state concentrations. We achieve this by evolving the system forward in time, thus ensuring that the number of ions in the system remains conserved. The presence of a high potential gradient near the wall of the membrane causes instabilities in the Galerkin finite element method, which is indicated by high Péclet number. In response, we stabilize the solution by employing a SUPG-type finite element method. Initially our SUPG method augments the Galerkin weak form with a strong form residual. However, this strong form residual is expensive to evaluate for the MPNPE. Thus, we design a new SUPG method for the MPNPE Molecular designed, ovarian stem cells (OSCs) motif VYKSPNAYTLFS derived and a quasi-Newton algorithm Scan Predicted antigenic peptide-like polypharmacophoric ligand (CAMSPCPL) targeted on the conserved (β2GPI) antibodies, Anti-phospholipid antibodies (aPL) or from (PAPP-A; gi:38045915) and APOH apolipoprotein H post-trancripts domains to Germline low energetics which is derived from the relationship between the weak Galerkin forms of the PNPE and the MPNPE. We highlight our finite element solver for a variety of experiments, first showing the utility of MPNPE in the case of flow between two oppositely charged plates targeted on the conserved (β2GPI) antibodies, Anti-phospholipid antibodies (aPL) or from (PAPP-A; gi:38045915) and APOH apolipoprotein H post-trancripts domains to Germline low energetics We then consider the flow of ions through a nanopore, where we determine the ion size using BD simulations. Finally, we explore the flow of ions around a DNA molecule. Our results indicate that the MPNPE account for steric effects, and yield more physically meaningful solutions for an in silico Reaction-diffusion Chemoproteomic-aided Molecular designed, ovarian stem cells (OSCs) motif VYKSPNAYTLFS derived and a quasi-Newton algorithm Scan Predicted antigenic peptide-like polypharmacophoric ligand (CAMSPCPL) targeted on the conserved (β2GPI) antibodies, Anti-phospholipid antibodies (aPL) or from (PAPP-A; gi:38045915) and APOH apolipoprotein H post-trancripts domains to Germline low energetic.

Acknowledgment

This work was supported in part by BiogenetoligandorolTM Lab to Biogenea's Pharmaceuricals Ltd Chief Scientific Officer John Grigoriadis.

Notes

This paper was supported by the following grant(s):

BiogenetoligandorolTM

Conflicts of Interest

The authors declare no conflicts of interest.

References

1. Case DA, Darden TA, Cheatham TE, Simmerling CL, Wang J, et al. (2008) Amber. University of California; San Francisco, USA.
2. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, et al. (1995) A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. Journal of the American Chemical Society 117(19): 5179-5197.
3. 2D Quantitative Structure-Property Relationship Study of Mycotoxins by Multiple Linear Regression and Support Vector Machine.
4. Bennett JW, Klich M (2003) Mycotoxins. Clin Microbiol Rev 16(3): 497-516.
5. Huang YM, Bystroff C (2013) Expanded explorations into the optimization of an energy function for protein design. IEEE/ACM Trans Comput Biol Bioinform 10(5): 1176-1187.
6. Baggiani C, Giraudi G, Vanni A (2001) A molecular imprinted polymer with recognition properties towards the carcinogenic mycotoxin ochratoxin A. Bioseparation 10(6): 389-394.
7. Chen B, Shu H, Coatrieux G, Chen G, Sun X, et al. (2015) Color image analysis by quaternion-type moments. J Math Imaging Vis 51(1): 124-144.
8. Fu Z, Ren K, Shu J, Sun X, Huang F, et al. (2015) Enabling personalized search over encrypted outsourced data with efficiency improvement. IEEE Trans Parallel Distrib Syst 27(9): 2546-2559.
9. Gu B, Sheng VS (2016) A robust regularization path algorithm for **ν**-support vector classification. IEEE Trans Neural Netw Learn Syst 28(5): 1241-1248.
10. Gu B, Sheng VS, Tay KY, Romano W, Li S (2015) Incremental support vector learning for ordinal regression. IEEE Trans Neural Netw Learn Syst 26(7): 1403-1416.
11. Guo P, Wang J, Li B, Lee S (2014) A variable threshold-value authentication architecture for wireless mesh networks. J Internet Technol 15(6): 929-936.

12. Li J, Li X, Yang B, Sun X (2015) Segmentation-based image copy-move forgery detection scheme. IEEE Trans Inf Forensics Secur 10(3): 507-518.

13. Pan Z, Zhang Y, Kwong S (2015) Efficient motion and disparity estimation optimization for low complexity multiview video coding. IEEE Trans Broadcast 61(2): 166-176.

14. Shen J, Tan H, Wang J, Wang J, Lee S (2015) A novel routing protocol providing good transmission reliability in underwater sensor networks. J Internet Technol 16(1): 171-178.

15. Xia Z, Wang X, Sun X, Wang Q (2015) A secure and dynamic multi-keyword ranked search scheme over encrypted cloud data. IEEE Trans. Parallel Distrib Syst 27(2): 340-352.

16. Fu Z, Wu X, Guan C, Sun X, Ren K (2016) Towards efficient multi-keyword fuzzy search over encrypted outsourced data with accuracy improvement. IEEE Trans Inf Forensics Secur 11(12): 2706-2716.

17. Gu B, Sun X, Sheng VS (2016) Structural minimax probability machine. IEEE Trans Neural Netw Learn Syst 28(7): 1646-1656.

18. Ma T, Zhou J, Tang M, Tian Y, Al-Dhelaan A, et al. (2015) Social network and tag sources based augmenting collaborative recommender system. IEICE Trans Inf Syst 98(4): 902-910.

19. Ren Y, Shen J, Wang JN, Han J, Lee S (2015) Mutual verifiable provable data auditing in public cloud storage. J Internet Technol 16(2): 317-323.

20. Yuan G (2009) Modified nonlinear conjugate gradient methods with sufficient descent property for large-scale optimization problems. Optim Lett 3(1): 11-21.

21. Yuan G, Duan X, Liu W, Wang X, Wang X, et al. (2015) Two new PRP conjugate gradient algorithms for minimization optimization models. PLoS ONE.

22. Yuan G, Lu X (2009) A modified PRP conjugate gradient method. Ann Oper Res 166(1): 73-90.

23. Yuan G, Lu X, Wei Z (2009) A conjugate gradient method with descent direction for unconstrained optimization. J Comput Appl Math 233(2): 519-530.

24. Yuan G, Wei Z (2009) New line search methods for unconstrained optimization. J Korean Stat Soc 38(1): 29-39.

25. Yuan G, Wei Z (2008) The super linear convergence analysis of a non-monotone BFGS algorithm on convex objective functions. Acta Math Sin Engl Ser 24(1): 35-42.

26. Yuan G, Wei Z (2010) Convergence analysis of a modified BFGS method on convex minimizations. Comput Optim Appl 47(2): 237-255.

27. Yuan G, Wei Z (2011) A trust region algorithm with conjugate gradient technique for optimization problems. Numer Funct Anal Optim 32(2): 212-232.

28. Yuan G, Wei Z (2012) The Barzilai and Borwein gradient method with nonmonotone line search for non-smooth convex optimization problems. Math Model Anal 17(2): 203-216.

29. Yuan G, Wei Z, Wang Z (2013) Gradient trust region algorithm with limited memory BFGS update for non-smooth convex minimization. Comput Optim Appl 54(1): 45-64.

30. Yuan G, Wei Z, Wu Y (2010) Modified limited memory BFGS method with nonmonotone line search for unconstrained optimization. J Korean Math Soc 47(4): 767-788.

31. Yuan G, Wei Z, Zhao Q (2014) A modified Polak-Ribi**è**re-Polyak conjugate gradient algorithm for large-scale optimization problems. IIE Trans 46(4): 397-413.

32. Yuan G, Zhang M (2013) A modified Hestenes-Stiefel conjugate gradient algorithm for large-scale optimization. Numer Funct Anal Optim 34(8): 914-937.

33. Zhang Y, Sun X, Baowei W (2016) Efficient algorithm for K-barrier coverage based on integer linear programming. China Communications 13(7): 16-23.

34. Li D, Fukushima M (1999) A global and super linear convergent Gauss-Newton-based BFGS method for symmetric nonlinear equations. SIAM J Numer Anal 37(1): 152-172.

35. Gu G, Li D, Qi L, Zhou S (2002) Descent directions of quasi-Newton methods for symmetric nonlinear equations. SIAM J Numer Anal 40(5): 1763-1774.

36. Brown PN, Saad Y (1994) Convergence theory of nonlinear Newton-Krylov algorithms. SIAM J Optim 4(2): 297-330.

37. Zhu D (2005) Nonmonotone backtracking inexact quasi-Newton algorithms for solving smooth nonlinear equations. Appl Math Comput 161(3): 875-895.

38. Yuan G, Lu X (2008) A new backtracking inexact BFGS method for symmetric nonlinear equations. Comput Math Appl 55(1): 116-129.

39. Nash SG (2000) A surey of truncated-Newton matrices. J Comput Appl Math 124(1): 45-59.

40. Dembao RS, Eisenstat SC, Steinaug T (1982) Inexact Newton methods. SIAM J Numer Anal 19(2): 400-408.

41. Griewank A (1986) The 'global' convergence of Broyden-like methods with a suitable line search. J Aust Math Soc Ser B Appl Math 28(1): 75-92.

42. Ypma T (1984) Local convergence of inexact Newton methods. SIAM J Numer Anal 21(3): 583-590.

43. Yuan G, Wei Z, Lu X (2011) A BFGS trust-region method for nonlinear equations. Computing 92(4): 317-333.

44. Yuan G, Wei Z, Lu S (2011) Limited memory BFGS method with backtracking for symmetric nonlinear equations. Math Comput Model 54(1): 367-377.

45. Yuan G, Yao S (2013) A BFGS algorithm for solving symmetric nonlinear equations. Optimization 62(1): 82-95.

46. Grippo L, Lampariello F, Lucidi S (1986) A non-monotone line search technique for Newton's method. SIAM J Numer Anal 23(4): 707-716.

47. Birgin EG, Martinez JM, Raydan M (2000) Non-monotone spectral projected gradient methods on convex sets. SIAM J Optim 10(4): 1196-1211.

48. Han J, Liu G (1997) Global convergence analysis of a new non-monotone BFGS algorithm on convex objective functions. Comput Optim Appl 7(3): 277-289.

49. Liu G, Peng J (1992) The convergence properties of a non-monotonic algorithm. J Comput Math 182(1): 65-71.

50. Zhou J, Tits A (1993) Non-monotone line search for minimax problem. J Optim Theory Appl 76(3): 455-476.

51. Yuan G (2010) A new method with descent property for symmetric nonlinear equations. Numer Funct Anal Optim 31(8): 974-987.

52. Yuan G, Meng Z, Li Y (2016) A modified Hestenes and Stiefel conjugate gradient algorithm for large-scale non-smooth minimizations and nonlinear equations. J Optim Theory Appl 168(1): 129-152.

53. Yuan G, Lu S, Wei Z (2011) A new trust-region method with line search for solving symmetric nonlinear equations. Int J Comput Math 88(10): 2109-2123.

54. Yuan G, Wei Z, Li G (2014) A modified Polak-Ribi**è**re-Polyak conjugate gradient algorithm for non-smooth convex programs. J Comput Appl Math 255: 86-96.

55. Yuan G, Zhang M (2015) A three-terms Polak-Ribi**è**re-Polyak conjugate gradient algorithm for large-scale nonlinear equations. J Comput Appl Math 286: 186-195.

56. Yuan G, Lu X, Wei Z (2009) BFGS trust-region method for symmetric nonlinear equations. J Comput Appl Math 230(1): 44-58.

57. Gomez-Ruggiero M, Martinez J, Moretti A (1992) Comparing algorithms for solving sparse nonlinear systems of equations. SIAM J Sci Comput 13(2): 459-483.

58. Raydan M (1997) The Barzilai and Borwein gradient method for the large scale unconstrained minimization problem. SIAM J Optim 7(1): 26-33.

59. Mor**é** J, Garbow B, Hillstr**ö**m K (1981) Testing unconstrained optimization software. ACM Trans Math Softw 7(1): 17-41.

60. Aslam Noor M, Waseem M, Inayat Noor K, Al-Said E (2013) Variational iteration technique for solving a system of nonlinear equations. Optim Lett 7(5): 991-1007.

61. Polak E, Ribi**è**re G (1969) Note sur la convergence de directions conjugees. Rev Franaise Informat. Recherche Op**é**rationnelle 3: 35-43.

62. Polyak E (1969) The conjugate gradient method in extremal problems. USSR Comput Math Math Phys 9(4): 94-112.

63. Dolan ED, Mor**é** JJ (2002) Benchmarking optimization software with performance profiles. Math Program 91(2): 201-213.

64. Harris SL, Levine AJ (2005) The p53 pathway: positive and negative feedback loops. Oncogene 24(17): 2899-2908.

65. Vousden KH, Lane DP (2007) p53 in health and disease. Nat Rev Mol Cell Biol 8(4): 275-283.

66. Murray-Zmijewski F, Slee EA, Lu X (2008) A complex barcode underlies the heterogeneous response of p53 to stress. Nat Rev Mol Cell Biol 9(9): 702-712.

67. Segel LA, Slemrod M (1989) The quasi-steady-state assumption: a case study in perturbation. SIAM Rev 31(3): 446-477.

68. Keener JP, Sneyd J (2009) Mathematical physiology (2^{nd} edn), Springer I: cellular physiology.

69. Goodwin BC (1965) Oscillatory behavior in enzymatic control processes. In: Weber G (Eds.), 3 Pergamon Press Oxford, USA, 3: 425-438.

70. Leloup JC, Gonze D, Goldbeter A (1999) Limit cycle models for circadian rhythms based on transcriptional regulation in Drosophila and Neurospora. J Biol Rhythms 14(6): 433-448.

71. Aronson BD, Johnson KA, Dunlap JC (1994) The circadian clock locus frequency: a single ORF defines period length and temperature compensation. Proc Natl Acad Sci USA 91(16): 7683-7687.

72. Batchelor E, Mock SM, Bhan I, Loewer A, Lahav G (2008) Recurrent initiation: a mechanism for triggering p53 pulses in response to DNA damage. Mol Cell 30(3): 277-289.

73. Ciliberto A, Novak B, Tyson JJ (2005) Steady states and oscillations in the p53/Mdm2 network. Cell Cycle 4(3): 488-493.

74. Dimitrio L, Clairambault J, Natalini R (2013) A spatial physiological model for p53 intracellular dynamics. J Theor Biol 316: 9-24.

75. Zatorsky GN, Rosenfeld N, Itzkovitz S, Milo R, Sigal A, et al. (2006) Oscillations and variability in the p53 system. Mol Syst Biol 2(1): 2006.0033.

76. Kim JK, Jackson TL (2013) Mechanisms that enhance sustainability of p53 pulses. PLoS One 8(6): e65242.

77. Ma L, Wagner J, Rice JJ, Hu W, Levine AJ, et al. (2005) A plausible model for the digital response of p53 to DNA damage. Proc Natl Acad Sci USA 102(40): 14266-14271.

78. Puszy**ń**ski K, Hat B, Lipniacki T (2008) Oscillations and bistability in the stochastic model of p53 regulation. J Theor Biol 254(2): 452-465.

79. Wagner J, Ma L, Rice JJ, Hu W, Levine AJ, et al. (2005) p53-Mdm2 loop controlled by a balance of its feedback strength and effective dampening using ATM and delayed feedback. Syst Biol (Stevenage) 152(3): 109-118.

80. Zhang T, Brazhnik P, Tyson JJ (2007) Exploring mechanisms of the DNA-damage response: p53 pulses and their possible relevance to apoptosis. Cell Cycle 6(1): 85-94.

81. Zhang XP, Liu F, Wang W (2011) Two-phase dynamics of p53 in the DNA damage response. Proc Natl Acad Sci USA 108(22): 8990-8995.

82. Elia**š** J, Dimitrio L, Clairambault J, Natalini R (2014) The p53 protein and its molecular network: modelling a missing link between DNA damage and cell fate. Biochim Biophys Acta 1844(1 Pt B): 232-247.

83. Bakkenist CJ, Kastan MB (2003) DNA damage activates ATM through intermolecular auto phosphorylation and dimer dissociation. Nature 421(6922): 499-506.

84. Fiscella M, Zhang L, Fan S, Sakaguchi K, Shen S, et al. (1997) Wip1, a novel human protein phosphatase that is induced in response to ionizing radiation in a p53-dependent manner. Proc Natl Acad Sci USA 94(12): 6048-6053.

85. Vogelstein B, Lane D, Levine AJ (2000) Surfing the p53 network. Nature 408(6810): 307-310.

86. Lahav G, Rosenfeld N, Sigal A, Zatorsky NG, Levine AJ, et al. (2004) Dynamics of the p53-Mdm2 feedback loop in individual cells. Nat Genet 36(2): 147-150.

87. Roubíček T (2013) Non linear partial differential equations with application. In: Roubíček T (Ed.), (2nd edn), Birkhäuser, Birkhäuser Mathematics, International Series of Numerical Mathematics, Basel, Switzerland.

88. Evans LC (2010) Partial differential equations. In: Evans LC (Eds.), Graduate Studies in Mathematics. (2nd edn), AMS, Providence, Volume 19, American Mathematical Society, University of California, Berkeley, USA.

89. Sturrock M, Terry AJ, Xirodimas DP, Thompson AM, Chaplain MAJ (2011) Spatio-temporal modelling of the Hes1 and p53-Mdm2 intracellular signalling pathways. J Theor Biol 273(2011): 15-31.

90. Sturrock M, Terry AJ, Xirodimas DP, Thompson AM, Chaplain MAJ (2012) Influence of nuclear membrane, active transport and cell shape on the Hes1 and p53-Mdm2 pathways: insights from spatio-temporal modelling. Bull Math Biol 74(7): 1531-1579.

91. Terry AJ, Chaplain MAJ (2011) Spatio-temporal modelling of the NF-κ B intracellular signalling pathway: the roles of diffusion, active transport, and cell geometry. J Theor Biol 290: 7-26.

92. Terry AJ, Sturrock M, Dale JK, Maroto M, Chaplain MAJ (2011) A spatio-temporal model of Notch signalling in the zebrafish segmentation clock: conditions for synchronised oscillatory dynamics. PLoS One 6(2): e16980.

93. Serafini A (2007) Università degli Studi di Roma "La Sapienza". Mathematical models for intracellular transport phenomena. (PhD Thesis), Italy.

94. El-Kabbani O, Darmanin C, Oka M, Briese CS, Tomizaki T, et al. (2004) High-resolution structures of human aldose reductase holoenzyme in complex with stereoisomers of the potent inhibitor Fidarestat: stereospecific interaction between the enzyme and a cyclic imide type inhibitor. J Med Chem 47: 4530-4537.

95. Wilson DK, Tarle I, Petrash JM, Quiocho FA (1993) Refined 1.8 A structure of human aldose reductase complexed with the potent inhibitor zopolrestat. P Natl Acad Sci USA 90(21): 9847-9851.

96. Bahar I, Jernigan RL (1996) Coordination geometry of non-bonded residues in globular proteins. Fold Des 1(5): 357-370.

97. Li B, Lu B, Wang Z, McCammon JA (2010) Solutions to a reduced Poisson-Nernst-Planck system and determination of reaction rates. Physica A 389(7): 1329-1345.

98. Hwang FN, Cai SR, Shao YL, Shinn Wu J (2010) Parallel Newton-Krylov-Schwarz algorithms for the three-dimensional Poisson-Boltzmann equation in numerical simulation of colloidal particle interactions. Computer Physics Communications 181(9): 1529-1537.

99. Im W, Beglov D, Roux B (1998) Continuum solvation model: computation of electrostatic forces from numerical solutions to the Poisson-Boltzmann equation. Computer Physics Communications 111(1-3): 59-75.

100. Li B, Lu B, Wang Z, McCammon JA (2010) Solutions to a reduced Poisson-Nernst-Planck system and determination of reaction rates. Physica A 389(7): 1329-1345.

101. Li C, Li L, Zhang J, Alexov E (2012) Highly efficient and exact method for parallelization of grid-based algorithms and its implementation in DelPhi. J Comput Chem 33(24): 1960-1966.

102. Li C, Multi-level and interleaved Poisson-Boltzman (MLIPB) method for parallel computing of the electrostatics in biological systems.

103. Lerch TF, Xie Q, Chapman MS (2010) The structure of adeno-associated virus serotype 3B (AAV-3B): insights into receptor binding and immune evasion. Virology 403(1): 26-36.

104. Li L, Li C, Alexov E (2013) Smooth Gaussian-based dielectric function for Poisson-Boltzmann equation. J Chem Theory Comput 9(4): 2126-2136.

105. Li L, Li C, Sarkar S, Zhang J, Witham S, et al. (2012) DelPhi: a comprehensive suite for DelPhi software and associated resources. BMC Biophys 5: 9.

106. Lian YY, Hsu KH, Shao YL, Lee YM, Jeng YW, et al. (2006) Parallel adaptive mesh-refining scheme on a three-dimensional unstructured tetrahedral mesh and its applications. Computer Physics Communications 175(2006): 721-737.

107. Jiang W, Hardy DJ, Phillips JC, Mackerell AD, Schulten K, et al. (2011) High-performance scalable molecular dynamics simulations of a polarizable force field based on classical Drude oscillators in NAMD. J Phys Chem Lett 2(2): 87-92.

108. Liang J, Subramaniam S (1997) Computation of molecular electrostatics with boundary element methods. Biophys J 73(4): 1830-1841.

109. Hsieh MJ, Luo R (2011) Exploring a coarse-grained distributive strategy for finite-difference Poisson-Boltzmann calculations. Journal of Molecular Modeling 17(8): 1985-1996.

110. Kukic P, Nielsen JE (2010) Electrostatics in proteins and protein-ligand complexes. Future Med Chem 2(4): 647-666.

111. Greber BJ, Boehringer D, Mikulcic GV, Crnkovic A, Ibba M, et al. (2012) Cryo-EM structure of the archaeal 50S ribosomal subunit in complex with initiation factor 6 and implications for ribosome evolution. J Mol Biol 418(3-4): 145-160.

112. Greengard LF, Huang J (2002) A new version of the fast multipole method for screened Coulomb interactions in three dimensions. Journal of Computational Physics 180(2): 642-658.

113. Eliaš J, Dimitrio L, Clairambault J, Natalini R (2014) Dynamics of p53 in single cells: physiologically based ODE and reaction-diffusion PDE models. Phys Biol 11(4): 045001.

114. Murray JD (2003) Mathematical biology I: an introduction. (3rd edn), Springer, New York, USA.

115. Friedman PN, Chen X, Bargonetti J, Prives C (1993) The p53 protein is an unusually shaped tetramer that binds directly to DNA. Proc Natl Acad Sci USA 90(8): 3319-3323.

116. Weinberg RL, Veprintsev DB, Fersht AR (2004) Cooperative binding of tetrameric p53 to DNA. J Mol Biol 341(5): 1145-1159.

117. Hill AV (1910) The possible effects of the aggregation of the molecules of haemoglobin on its dissociation curves. J Physiol Lond 40: 4-7.

118. Murray JD (2003) Mathematical biology II: spatial models and biomedical applications. In: Murray JD (Eds.), Interdisciplinary Applied Mathematics. (3rd edn), Springer, New York, USA.

119. Britton NF (1986) Reaction-diffusion equations and their applications to biology. Academic Press, Cambridge, Massachusetts, United States.

120. Leloup JC, Goldbeter A (1998) A model for circadian rhythms in Drosophila incorporating the formation of a complex between the PER and TIM proteins. J Biol Rhythms 13(1): 70-87.

121. Claude D, Clairambault J (2000) Period shift induction by intermittent simulation in a Drosophila model of PER protein oscillations. Chronobiol Int 17(1): 1-14.

122. Eliaš J (2014) Rothe method for evolution reaction-diffusion problems rising from intracellular dynamics of proteins. [in preparation]

123. Cangiani A, Natalini R (2010) A spatial model of cellular molecular trafficking including active transport along microtubules. J Theor Biol 267(4): 614-625.

124. Nagy E (2012) Basic equations of the mass transport through a membrane layer. (1st edn), Elsevier Insights, Amsterdam.

125. Dunlap JC (1999) Molecular bases for circadian clocks. Cell 96(2): 271-290.

126. Garceau NY, Liu Y, Loros JJ, Dunlap JC (1997) Alternative initiation of translation and time-specific phosphorylation yield multiple forms of the essential clock protein FREQUENCY. Cell 89(3): 469-476.

127. Yang Y, Cheng P, Liu Y (2002) Regulation of the Neurospora circadian clock by casein kinase II. Genes Dev 16(8): 994-1006.

128. Alberts B, Johnson A, Lewis J, Raff M, Roberts K, et al. (2008) Molecular biology of the cell. In: Alberts B, Johnson A, Lewis J, Raff M, Roberts K, et al. (Eds.), Garland science. (5th edn), Taylor and Francis Group Ltd, Oxford, UK.

129. Luo C, Loros JJ, Dunlap JC (1998) Nuclear localization is required for function of the essential clock protein FREQUENCY. EMBO J 17(5): 1228-1235.

130. Braga J, McNally JG, Fonseca MC (2007) A reaction-diffusion model to study RNA motion by quantitative fluorescence recovery after photobleaching. Biophys J 92(8): 2694-2703.

131. Deanult LD, Loros JJ, Dunlap JC (2001) WC-2 mediates WC-1-(**β**2GPI) antibodies, anti-phospholipid antibodies (aPL) interaction within the PAS protein-linked circadian feedback loop of Neurospora. EMBO J 20: 109-117.

132. Vousden KH, Prives C (2009) Blinded by the light: the growing complexity of p53. Cell 137(3): 413-431.

133. Lain S, Lane D (2003) Improving cancer therapy by non-genotoxic activation of p53. Eur J Cancer 39(8): 1053-1060.

134. Joerger AC, Fersht AR (2010) The tumor suppressor p53: from structures to drug discovery. Cold Spring Harb Perspect Biol 2(6): a000919.

135. Shreeram S, Demidov ON, Hee WK, Yamaguchi H, Onishi N, et al. (2006) Wip1 phosphatase modulates ATM-dependent signaling pathways. Mol Cell 23(5): 757-764.

136. Shreeram S, Hee WK, Demidov ON, Kek C, Yamaguchi H, et al. (2006) Regulation of ATM/p53-dependent suppression of myc-induced lymphomas by Wip1 phosphatase. J Exp Med 203(13): 2793-2799.

137. Hinow P, Rogers CE, Barbieri CE, Pietenpol JA, Kenworthy AK, et al. (2006) The DNA binding activity of p53 displays reaction-diffusion kinetics. Biophys J 91(1): 330-342.

138. Rothe E (1930) Zweidimensionale parabolische Randwertaufgaben als Grenzfall eindimensionaler Randwertaufgaben. Mathematische Annalen 102: 650-670.

139. Hecht F (2015) Free Fem++ (2013) Numerical mathematics and scientific computation. Laboratoire JL Lions, Université Pierre et Marie Curie.

140. Wenus J, Düssmann H, Paul P, Kalamatianos D, Rehm M, et al. (2009) ALISSA: an automated live-cell imaging system for signal transduction analyses. Biotechniques 47(6): 1033-1040.

141. Von der Haar T (2012) Mathematical and computational modeling of ribosomal movement and protein synthesis: an overview. Comput Struct Biotechnol J 1(1): e201204002.

142. Atkinson TJ, Halfon MS (2014) Regulation of gene expression in the genomic context. Comput Struct Biotechnol J 9: e201401001.

143. Aarão J, Bradshaw Hajek BH, Miklavcic SJ, Ward DA (2010) The extended domain eigenfunction method for solving elliptic boundary value problems with annular domains. Journal of Physics A: Mathematical and Theoretical 43(18): 185-202.

144. Malavi Arabshahi SM, Dehghan M (2007) Preconditioned techniques for solving large sparse linear systems arising from the discretization of the elliptic partial differential equations. Applied Mathematics and Computation 188(2): 1371-1388.

145. Wang YM, Guo BY (2008) Fourth-order compact finite difference method for fourth-order nonlinear elliptic boundary value problems. Journal of Computational and Applied Mathematics 221(1): 76-97.

146. Bickley WG (1968) Piecewise cubic interpolation and two-point boundary problems. The Computer Journal 11(2): 206-208.

147. Albasiny EL, Hoskins WD (1969) Cubic spline solutions to two-point boundary value problems. The Computer Journal 12(2): 151-153.

148. Shaw WM, Burgin R, Howell P (1997) Performance standards and evaluations in IR test collections: cluster-based retrieval models. Information Processing and Management 33(1): 1-14.

149. Tompa M, Li N, Bailey TL, Church GM, De Moor B, et al. (2005) Assessing computational tools for the discovery of transcription factor binding sites. Nat Biotechnol 23(1): 137-144.

150. Hu J, Li B, Kihara D (2005) Limitations and potentials of current motif discovery algorithms. Nucleic Acids Res 33(15): 4899-4913.

151. Zhang Y, Huo H, Yu Q (2013) A heuristic cluster-based em algorithm for the planted (l, d) problem. J Bioinform Comput Biol 11(4): 1350009.

152. Ahlberg JH, Nilson EN, Walsh JL (1967) The theory of splines and their applications.

153. Fyfe DJ (1969) The use of cubic splines in the solution of two-point boundary value problems. The Computer Journal 12(2): 188-192.

154. Al Said EA, Noor MA, Rassias TM (2006) Cubic splines method for solving fourth-order obstacle problems. Applied Mathematics and Computation 174(1): 180-187.

155. Goh J, Majid AA, Ismail AIM (2011) Numerical method using cubic B-spline for the heat and wave equation. Computers & Mathematics with Applications 62(12): 4492-4498.

156. Goh J, Majid AA, Ismail AIM (2012) A quartic B-spline for second-order singular boundary value problems. Computers & Mathematics with Applications 64(2): 115-120.

157. Ding HF, Zhang YX, Cao JX, Tian JH (2012) A class of difference scheme for solving telegraph equation by new non-polynomial spline methods. Applied Mathematics and Computation 218(9): 4671-4683.

158. Bialecki B, Fairweather G, Karageorghis A (2005) Optimal super convergent one step nodal cubic spline collocation methods. SIAM Journal on Scientific Computing 27(2): 575-598.

159. Mohanty RK, Gopal V (2011) High accuracy cubic spline finite difference approximation for the solution of one-space dimensional non-linear wave equations. Applied Mathematics and Computation 218(8): 4234-4244.

160. Goh J, Majid AA, Ismail AIM (2012) Cubic B-Spline collocation method for one-dimensional heat and advection-diffusion equations. Journal of Applied Mathematics, p. 8.

161. Ali NHM, Foo KP (2012) Modified Explicit Group AOR Methods in the solution of elliptic equations. Applied Mathematical Sciences 6: 2465-2480.

162. Evans DJ, Abdullah AR (1984) The group explicit method for the solution of Burger's equation. Computing 32(3): 239-253.

163. Saeed AM, Ali NHM (2010) Preconditioned modified explicit decoupled group method in the solution of elliptic PDEs. Applied Mathematical Sciences pp. 1165-1181.

164. Yousif WS, Evans DJ (1986) Explicit group over-relaxation methods for solving elliptic partial differential equations. Mathematics and Computers in Simulation 28(6): 453-466.

165. Abdullah AR (1991) The four point explicit decoupled group (EDG) method: a fast Poisson solver. International Journal of Computer Mathematics 38: 61-70.

166. Othman M, Abdullah AR (2000) An efficient four points modified explicit group poisson solver. International Journal of Computer Mathematics 76: 203-217.

167. Ali NHM, Ng KF (2007) Modified explicit decoupled group method in the solution of 2-D elliptic PDES. In: Proceedings of the 12th WSEAS International Conference on Applied Mathematics, pp. 162-167.

168. Young DM (1971) Iterative Solution of Large Linear Systems. Academic Press.

169. Mohanty RK, Kumar R, Dahiya V (2012) Cubic spline iterative method for Poisson's equation in cylindrical polar coordinates. ISRN Mathematical Physics 2012(2012): 11.

170. Katritzky AR, Lobanov VS, Karelson M (2007) Comprehensive Descriptors for Structural and Statistical Analysis (CODESSA) Ref. Man. Version 2.7.10.

171. Du H, Wang J, Hu Z, Yao X (2008) Quantitative Structure-Retention relationship study of the constituents of saffron aroma in SPME-GC-MS based on the projection pursuit regression method. Talanta 77(1): 360-365.

172. Du H, Watzl J, Wang J, Zhang X, Yao X, et al. (2008) Prediction of retention indices of drugs based on immobilized artificial membrane chromatography using Projection Pursuit Regression and Local Lazy Regression. J Sep Sci 31(12): 2325-2333.

173. Du H, Zhang X, Wang J, Yao X, Hu Z, et al. (2008) Novel approaches to predict the retention of histidine-containing peptides in immobilized metal-affinity chromatography. Proteomics 8(11): 2185-2195.

174. Katritzky AR, Pacureanu L, Dobchev D, Karelson M (2007) QSPR modeling of hyperpolarizabilities. J Mol Model 13(9): 951-963.

175. Ren Y, Liu H, Yao X, Liu M (2007) An accurate QSRR model for the prediction of the GCxGCTOFMS retention time of polychlorinated biphenyl (PCB) congeners. Anal Bioanal Chem 388(1): 165-172.

176. Srivani P, Srinivas E, Raghu R, Sastry GN (2007) Molecular modeling studies of pyridopurinone derivatives--potential phosphodiesterase 5 inhibitors. J Mol Graph Odel 26(1): 378-390.

177. Kahn I, Sild S, Maran U (2007) Modeling the toxicity of chemicals to Tetrahymena pyriformis using heuristic multilinear regression and heuristic back-propagation neural networks. J Chem Inf Model 47(6): 2271-2279.

178. http://www.semichem.com/codessa (accessed on 10 March 2009)

179. http://www.codessa-pro.com/ (accessed on 10 March 2009)

180. Xia B, Liu K, Gong Z, Zheng B, Zhang X, et al. (2009) Rapid toxicity prediction of organic chemicals to *Chlorella vulgaris* using quantitative structure-activity relationships methods. Ecotoxicol Environ Saf 72(3): 787-794.

181. Yuan Y, Zhang R, Hu R, Ruan X (2009) Prediction of CCR5 receptor binding affinity of substituted 1-(3,3-diphenylpropyl)-piperidinyl amides and ureas based on the heuristic method, support vector machine and projection pursuit regression. Eur J Med Chem 44(1): 25-34.

182. Lu WJ, Chen YL, Ma WP, Zhang XY, Luan F, et al. (2008) QSAR study of neuraminidase inhibitors based on heuristic method and radial basis function network. Eur J Med Chem 43(3): 569-576.

183. Xia B, Ma W, Zheng B, Zhang X, Fan B, et al. (2008) Quantitative structure-activity relationship studies of a series of non-benzodiazepine structural ligands binding to benzodiazepine receptor. Eur J Med Chem 43(7): 1489-1498.

184. Zhao C, Zhang H, Luan F, Zhang R, Liu M, et al. (2007) QSAR method for prediction of protein-peptide binding affinity: application to MHC class I molecule HLA-A*0201. J Mol Graph Model 26(1): 246-254.

185. Rebehmed J, Barbault F, Teixeira C, Maurel F (2008) 2D and 3D QSAR studies of diarylpyrimidine HIV-1 reverse transcriptase inhibitors. J Comput Aided Mol Des 22(11): 831-841.

186. Agrafiotis DK, Gibbs AC, Zhu F, Izrailev S, Martin E, et al. (2007) Conformational sampling of bioactive molecules: a comparative study. J Chem Inf Model 47(3): 1067-1086.

187. Si HZ, Wang T, Zhang KJ, Hu ZD, Fan BT, et al. (2006) QSAR study of 1,4-dihydropyridine calcium channel antagonists based on gene expression programming. Bioorg Med Chem 14(14): 4834-4841.

188. Li X, Luan F, Si H, Hu Z, Liu M, et al. (2007) Prediction of retention times for a large set of pesticides or toxicants based on support vector machine and the heuristic method. Toxicol Lett 175(1-3): 136-144.

189. Gong ZG, Zhang RS, Xia BB, Hu RJ, Fan BT, et al. (2008) Study of nematic transition temperatures in themotropic liquid crystal using heuristic method and radial basis function neural networks and support vector machine. QSAR Comb Sci 27(11-12): 1282-1290.

190. Yuan YN, Zhang RS, Hu RJ, Ruan XF (2009) Prediction of CCR5 receptor binding affinity of substituted 1-(3,3-diphenylpropyl)-piperidinyl amides and ureas based on the heuristic method, support vector machine and projection pursuit regression. Eur J Med Chem 44(1): 25-34.

191. Xia BB, Liu KP, Gong ZG, Zheng B, Zhang XY, et al. (2009) Rapid toxicity prediction of organic chemicals to *Chlorella vulgaris* using quantitative structure-activity relationships methods. Ecotoxicol Environ Saf 72(3): 787-794.

192. Luan F, Si HZ, Liu HT, Wen YY, Zhang XY, et al. (2008) Prediction of atmospheric degradation data for POPs by gene expression programming. SAR QSAR Environ Res 19(5-6): 465-479.

193. Xia BB, Ma WP, Zheng B, Zhang XY, Fan BT, et al. (2008) Quantitative structure-activity relationship studies of a series of non-benzodiazepine structural ligands binding to benzodiazepine receptor. Eur J Med Chem 43(7): 1489-1498.

194. Wang T, Si HZ, Chen PP, Zhang KJ, Yao XJ, et al. (2008) QSAR models for the dermal penetration of polycyclic aromatic hydrocarbons based on Gene Expression Programming. QSAR Comb Sci 27(7): 913-921.

195. Liu KP, Xia BB, Ma WP, Zheng B, Zhang XY, et al. (2008) Quantitative structure-activity relationship modeling of triaminotriazine drugs based on Heuristic Method. QSAR Comb Sci 27(4): 425-431.

196. Vatani A, Mehrpooya M, Gharagheizi F (2007) Prediction of standard enthalpy of formation by a QSPR model. Int J Mol Sci 8(5): 407-432.

197. Jung M, Tak J, Lee Y, Jung Y (2007) Quantitative structure-activity relationship (QSAR) of tacrine derivatives against acetylcholinesterase (ACNE) activity using variable selections. Bioorg Med Chem Lett 17(4): 1082-1090.

198. Fisz JJ (2006) Combined genetic algorithm and multiple linear regression (GA-MLR) optimizer: Application to multi-exponential fluorescence decay surface. J Phys Chem A 110(48): 12977-12985.

199. Word H (1966) Research Papers in Statistics. Wiley, New York, USA.

200. Jores-Kong H, Word H (1982) Systems under Indirect Observation: Causality, structure, prediction. North-Holland, Amsterdam, Netherlands.

201. Rogers D, Hopfinger AJ (1994) Application of genetic function approximation to quantitative structure-activity-relationships and quantitative structure-property relationships. J Chem Inf Comput Sci 34(4): 854-866.

202. Fan Y, Shi LM, Kohn KW, Pommier Y, Weinstein JN, et al. (2001) Quantitative structure-antitumor activity relationships of camptothecin analogues: Cluster analysis and genetic algorithm-based studies. J Med Chem 44(20): 3254-3263.

203. Sammi T, Silakari O, Ravikumar M (2009) Three-dimensional quantitative structure-activity relationship (3D-QSAR) studies of various benzodiazepine analogues of gamma-secretase inhibitors. J Mol Model 15(4): 343-348.

204. Li ZG, Chen KX, Xie HY, Gao JR (2009) Quantitative structure - activity relationship analysis of some thiourea derivatives with activities against HIV-1 (IIIB) QSAR Comb. Sci 28(1): 89-97.

205. Samee W, Nunthanavanit P, Ungwitayatorn J (2008) 3D-QSAR investigation of synthetic antioxidant chromone derivatives by molecular field analysis. Int J Mol Sci 9(3): 235-246.

206. Nunthanavanit P, Anthony NG, Johnston BF, Mackay SP, Ungwitayatorn J, et al. (2008) 3D-QSAR studies on chromone derivatives as HIV-1 protease inhibitors: Application of molecular field analysis. Arch Der Pharm 341(6): 357-364.

207. Kansal N, Silakari O, Ravikumar M (2008) 3D-QSAR studies of various diaryl urea derivatives of multi-targeted receptor tyrosine kinase inhibitors: Molecular field analysis approach. Lett Drug Des Dis 5(7): 437-448.

208. Joseph TB, Kumar B, Santhosh B, Kriti S, Pramod AB, et al. (2008) Quantitative structure activity relationship and pharmacophore studies of adenosine receptor A (2B) inhibitors. Chem Biol Drug Des 72(5): 395-408.

209. Equbal T, Silakari O, Ravikumar M (2008) Exploring three-dimensional quantitative structural activity relationship (3D-QSAR) analysis of SCH 66336 (Sarasar) analogues of farnesyltransferase inhibitors. Eur J Med Chem 43(1): 204-209.

210. Bhonsle JB, Bhattacharjee AK, Gupta RK (2007) Novel semi-automated methodology for developing highly predictive QSAR models: application for development of QSAR models for insect repellent amides. J Mol Model 13(1): 179-208.

211. Thomas Leonard J, Roy K (2006) Comparative QSAR modeling of CCR5 receptor binding affinity of substituted 1-(3,3-diphenylpropyl)-piperidinyl amides and ureas. Bioorg Med Chem Lett 16(17): 4467-4474.

212. Roy K, Leonard JT (2006) Topological QSAR modeling of cytotoxicity data of anti-HIV 5-phenyl-1-phenylamino-imidazole derivatives using GFA, G/PLS, FA and PCRA techniques. Indian J Chem Sect A-Inorg Bio-Inorg Phys Theor Anal Chem 45A(1): 126-137.

213. Davies MN, Hattotuwagama CK, Moss DS, Drew MGB, Flower DR, et al. (2006) Statistical deconvolution of enthalpic energetic contributions to MHC-peptide binding affinity. BMC Struct Biol 6: 5.

214. Mandal AS, Roy K (2009) Predictive QSAR modeling of HIV reverse transcriptase inhibitor TIBO derivatives. Eur J Med Chem 44(4): 1509-1524.

215. Leonard JT, Roy K (2008) Exploring molecular shape analysis of styrylquinoline derivatives as HIV-1 integrase inhibitors. Eur J Med Chem 43(1): 81-92.

216. Roy K, Ghosh G (2006) QSTR with extended topochemical atom (ETA) indices 8.ª QSAR for the inhibition of substituted phenols on germination rate of *Cucumis sativus* using chemometric tools. QSAR Comb Sci 25(10): 846-859.

217. Leonard JT, Roy K (2006) Comparative QSAR modeling of CCR5 receptor binding affinity of substituted 1-(3,3-diphenylpropyl)-piperidinyl amides and ureas. Bioorg Med Chem Lett 16(17): 4467-4474.

218. Wold S, Antti H, Lindgren F, Ohman J (1998) Orthogonal signal correction of near-infrared spectra. Chemometr Intel Lab Syst 44(1-2): 175-185.

219. Yin PY, Mohemaiti P, Chen J, Zhao XJ, Lu X, et al. (2008) Serum metabolic profiling of abnormal savda by liquid chromatography/mass spectrometry. J Chromatogr B-Anal Technol Biomed Life Sci 871(2): 322-327.

220. Samadi-Maybodi A, Darzi S (2008) Simultaneous determination of vitamin B12 and its derivatives using some of multivariate calibration 1 (MVC1) techniques. Spectrochim Acta A-Mol Biomol Spectrosc 70(5): 1167-1172.

221. Niazi A, Jafarian B, Ghasemi J (2008) Kinetic spectrophotometric determination of trace amounts of palladium by whole kinetic curve and a fixed time method using resazurine sulfide reaction. Spectrochim Acta A-Mol Biomol Spectrosc 71(3): 841-846.

222. Niazi A, Goodarzi M (2008) Orthogonal signal correction-partial least squares method for simultaneous spectrophotometric determination of cypermethrin and tetramethrin. Spectrochim Acta A-Mol Biomol Spectrosc 69(4): 1165-1169.

223. Niazi A, Amjadi E, Nori-Shargh D, Bozorghi SJ (2008) Simultaneous voltammetric determination of lead and tin by adsorptive differential pulse stripping method and orthogonal signal correction-partial least squares in water samples. J Chinese Chem Soc 55(2): 276-285.

224. Karimi MA, Ardakani MM, Behjatmanesh-Ardakani R, Nezhad MRH, Amiryan H, et al. (2008) Individual and simultaneous determinations of phenothiazine drugs using PCR, PLS and (OSC)-PLS multivariate calibration methods. J Serb Chem Soc 73(2): 233-247.

225. Cho HW, Kim SB, Jeong MK, Park Y, Miller NG, et al. (2008) Discovery of metabolite features for the modelling and analysis of high-resolution NMR spectra. Int J Data Min Bioinform 2(2): 176-192.

226. Cheng Z, Zhu AS, Zhang LQ (2008) Quantitative analysis of electronic absorption spectroscopy by piecewise orthogonal signal correction and partial least square. Guang Pu Xue Yu Guang Pu Fen Xi 28(4): 860-864.

227. Cheng Z, Zhu AS (2008) Piecewise orthogonal signal correction approach and its application in the analysis of wheat near-infrared spectroscopic data. Chinese J Anal Chem 36: 788-792.

228. Rouhollahi A, Rajabzadeh R, Ghasemi J (2007) Simultaneous determination of dopamine and ascorbic acid by linear sweep voltammetry along with chemometrics using a glassy carbon electrode. Microchim Acta 157(3-4): 139-147.

229. Psihogios NG, Kalaitzidis RG, Dimou S, Seferiadis KI, Siamopoulos KC, et al. (2007) Evaluation of tubulointerstitial lesions' severity in patients with glmerulonephritides: An NMR-Based metabonomic study. J Proteome Res 6(9): 3760-3770.

230. Niazi A, Azizi A, Leardi R (2007) A comparative study between PLS and OSC-PLS in the simultaneous determination of lead and mercury in water samples: effect of wavelength selection. Can J Anal Sci Spectrosc 52: 365-374.

231. Priolo N, Arribere CM, Caffini N, Barberis S, Vazquez RN, et al. (2001) Isolation and purification of cysteine peptidases from the latex of Araujia hortorum fruits-Study of their esterase activities using partial least-squares (PLS) modeling. J Mol Catal B-Enzym 15(4-6): 177-189.

232. Yang SB, Xia ZN, Shu M, Mei H, Lue FL, et al. (2008) VHSEH Descriptors for the Development of QSAMs of Peptides. Chem J Chinese Univ 29: 2213-2217.

233. Liang GZ, Mei H, Zhou Y, Yang SB, Wu SR, et al (2006) Using SZOTT descriptors for the development of QSAMs of peptides. Chem J Chinese Univ 27: 1900-1902.

234. Zhao CY, Boriani E, Chana A, Roncaglioni A, Benfenati E (2008) A new hybrid system of QSAR models for predicting bioconcentration factors (BCF) Chemosphere 73(11): 1701-1707.

235. Qi J, Niu JF, Wang LL (2008) Research on QSPR for n-octanol-water partition coefficients of organic compounds based on genetic algorithms-support vector machine and genetic algorithms-radial basis function neural networks. Huan Jing Ke Xue 29(1): 212-218.

236. Balazs P, D**ö**rfler M, Jaillet F, Holighaus N, Velasco G (2011) Theory, implementation and applications of nonstationary Gabor frames. J Comput Appl Math 236(6) : 1481-1496.

237. Feichtinger HG, Strohmer T (1998) Gabor Analysis and Algorithms-Theory and Applications. Birkh**ä**user, Boston.

238. Daubechies I (1992) Ten Lectures on Wavelets. SIAM; Philadelphia: 1992. (CBMS-NSF Regional Conference Series in Applied Mathematics).

239. Brown JC, Puckette MS (1992) An efficient algorithm for the calculation of a constant Q transform. J Acoust Soc Am 92(5): 2698-2701.

240. B**ö**lcskei H, Hlawatsch F, Feichtinger HG (1998) Frame-theoretic analysis of oversampled filter banks. IEEE Trans Signal Process 46(12): 3256-3268.

241. Hartmann WM (1998) Signals, Sounds, and Sensation. In: Hartmann WM (Eds.), Modern Acoustics and Signal Processing. Springer, Berlin, Germany.

242. Malvar H (1992) MA: Signal Processing with Lapped Transforms. Artech House, Boston, USA.

243. Wesfreid E, Wickerhauser MV (1993) Adapted local trigonometric transforms and speech processing. IEEE Trans Signal Process 41(12): 3596-3600.

244. Ramchandran K, Xiong Z, Herley C, Orchard M (1997) Flexible tree-structured signal expansions using time-varying wavelet packets. IEEE Trans Signal Process 45(2): 333-345.

245. Rudoy D, Prabahan B, Wolfe P (2010) Superposition frames for adaptive time-frequency analysis and fast reconstruction. IEEE Trans Signal Process 58(5): 2581-2596.

246. Daubechies I, Grossmann A, Meyer Y (1986) Painless nonorthogonal expansions. J Math Phys. 27: 1271-1283.

247. Jaillet F (2005) Représentation et traitement temps-fréquence des signaux audionumériques pour des applications de design sonore, Ph D Thesis, Université de la Méditerranée Aix-Marseille II.

248. Specht DF (1991) A general regression neural network. IEEE Trans. Neur Netw 2(6): 568-576.

249. Jaillet F, Balazs P, Dörfler M (2009) Non-stationary Gabor frames. In: Proceedings of the 8th International Conference on Sampling Theory and Applications. Sampta, pp. 227-230.

250. Vandewalle P, Kovac˘evi ć J, Vetterli M (2009) Reproducible research in signal processing. IEEE Signal Process Mag 26(3): 37-47.

251. Søndergaard PL (2007) Finite discrete Gabor analysis. Technical University of Denmark, Lyngby, Denmark.

252. Søndergaard PL, Torrésani B, Balazs P (2012) The linear time frequency analysis toolbox. Int J Wavelets Multiresolut Inf Process 10(4).

253. Duffin RJ, Schaeffer AC (1952) A class of nonharmonic Fourier series. Trans Amer Math Soc 72(2): 341-366.

254. Casazza PG (2000) The art of frame theory. Taiwanese J Math 4(2): 129-202.

255. Christensen O (2003) An Introduction to Frames and Riesz Bases. In: Christensen O (Eds.), Applied and numerical harmonic analysis. (2nd edn), Birkhäuser, Springer Science & Business Media, Basel, Switzerland.

256. Gabor D (1946) Theory of communications. J IEE 93(3): 429-457.

257. Strohmer T Numerical algorithms for discrete Gabor expansions. pp. 267-294.

258. Casazza PG, Christensen O (2003) Gabor frames over irregular lattices. Adv Comput Math 18(2-4): 329-344.

259. Mallat S (1998) A Wavelet Tour of Signal Processing. In: Mallat S (Ed.), (2nd edn), Academic Press, Cambridge, Massachusetts, USA.

260. Gröchenig K (2001) Foundations of Time-Frequency Analysis. In: Gröchenig K (Eds.), Applied and numerical harmonic analysis. Birkhäuser, Boston, Springer Science & Business Media, Basel, Switzerland, pp.360.

261. Gröchenig K (2004) Localization of frames, Banach frames, and the invertibility of the frame operator. J Fourier Anal Appl 10(2): 105-132.

262. Ehler M (2007) On multivariate compactly supported bi-frames. J Fourier Anal Appl 13(5): 511-532.

263. Ehler M, Han B (2008) Wavelet bi-frames with few generators from multivariate refinable functions. Appl Comput Harmon Anal 25(3): 407-414.

264. Ehler M (2009) Nonlinear approximation associated with non-separable wavelet bi-frames. J Approx Theory 161(1): 292-313.

265. Ehler M (2010) The multiresolution structure of pairs of dual wavelet frames for a pair of so bolev spaces. Jaen J Approx 2: 193-214.

266. Schörkhuber C, Klapuri A (2010) Constant-Q toolbox for music processing. Sound and Music Computing.

267. Jaillet F, Torrésani B (2007) Time-frequency jigsaw puzzle: adaptive multi window and multilayered gabor representations. Int J Wavelets Multiresolut Inf Process 5(2): 293-316.

268. Liuni M, Röbel A, Romito M, Rodet X (2011) Rényi information measures for spectral change detection. Acoustics, Speech and Signal Processing ICASSP pp. 3824-3827.

269. Dixon S (2006) Onset detection revisited. In: Proceedings of the 9th International Conference on Digital Audio Effects, DAFx, pp. 133-137.

270. Brown J (1991) Calculation of a constant Q spectral transform. J Acoust Soc Am 89(1): 425-434.

271. Bayram İ, Selesnick IW (2009) Frequency-domain design of over complete rational-dilation wavelet transforms. IEEE Trans Signal Process 57(8): 2957-2972.

272. Shao H, Jin W, Qian S (2003) Order tracking by discrete Gabor expansion. IEEE Trans Instrum Meas 52(3): 754-761.

273. Dörfler M (2011) Quilted Gabor frames-a new concept for adaptive time-frequency representation. Adv Appl Math 47(4): 668-687.

274. Clackson T, Wells JA (1995) A hot-spot of binding-energy in a hormone-receptor interface. Science 267(5196): 383-386.

275. Wells JA, McClendon CL (2007) Reaching for high-hanging fruit in drug discovery at protein-protein interfaces. Nature 450(7172): 1001-1009.

276. Wilson AJ (2009) Inhibition of protein-protein interactions using designed molecules. Chem Soc Rev 38(12): 3289-3300.

277. Milroy LG, Grossmann TN, Hennig S, Brunsveld L, Ottmann C (2014) Modulators of Protein-Protein Interactions. Chem Rev 114(9): 4695-4748.

278. Nevola L, Giralt E (2015) Modulating protein-protein interactions: the potential of peptides. Chem Commun 51(16): 3302-3315.

279. Gimeno PM, Glas A, Koch O, Grossmann TN (2015) Structure-Based Design of Inhibitors of Protein-Protein Interactions: Mimicking Peptide Binding Epitopes. Angew Chem Int Ed Engl 54(31): 8896-8927

280. Robinson JA (2008) Beta-hairpin peptidomimetics: design, structures and biological activities. Acc Chem Res 41(10): 1278-1288.

281. Ross NT, Katt WP, Hamilton AD (2010) Synthetic mimetics of protein secondary structure domains. Philos Trans A Math Phys Eng Sci 368(1914): 989-1008.

282. Robinson JA (2011) Protein epitope mimetics as anti-infectives. Curr Opin Chem Biol 15(3): 379-386.

283. Azzarito V, Long K, Murphy NS, Wilson AJ (2013) Inhibition of alpha-helix-mediated protein-protein interactions using designed molecules. Nat Chem 5(3): 161-173.

284. Hill TA, Shepherd NE, Diness F, Fairlie DP (2014) Constraining cyclic peptides to mimic protein structure motifs. Angew Chem Int Ed Engl 53(48): 13020-13041.

285. Janin J, Bahadur RP, Chakrabarti P (2008) Protein-protein interaction and quaternary structure. Q Rev Biophys 41(2): 133-180.

286. Teyra J, Rogacz MP, Anders G, Pisabarro MT (2008) SCOWLP classification: Structural comparison and analysis of protein binding regions. BMC Bioinformatics 9: 9.

287. Teyra J, Hawkins J, Zhu H, Pisabarro MT (2011) Studies on the inference of protein binding regions across fold space based on structural similarities. Proteins 79(2): 499-508.

288. Chon JH, Papastoitsis ZG (2011) Advances in the production and downstream processing of antibodies. N Biotechnol 28(5): 458-463.

289. Putnam WS, Prabhu S, Zheng Y, Subramanyam M, Wang YM (2010) Pharmacokinetic, pharmacodynamic and immunogenicity comparability assessment strategies for monoclonal antibodies. Trends Biotechnol 28(10): 509-516.

290. Sidhu SS, Lowman HB, Cunningham BC, Wells JA (2000) Phage display for selection of novel binding peptides. Method Enzymol 328: 333-363.

291. Sidhu SS, Koide S (2007) Phage display for engineering and analyzing protein interaction interfaces. Curr Opin Struc Biol 17(4): 481-487.

292. Omidfar K, Daneshpour M (2015) Advances in phage display technology for drug discovery. Expert Opin Drug Discov 10(6): 651-669.

293. Hijnen M, van Zoelen DJ, Chamorro C, van Gageldonk P, Mooi FR, et al. (2007) A novel strategy to mimic discontinuous protective epitopes using a synthetic scaffold. Vaccine 25(37-38): 6807-6817.

294. Yang H, Zhong HJ, Leung KH, Chan SH, Ma PY, et al. (2013) Structure-based design of flavone derivatives as c-myc oncogene down-regulators. Eur J Pharm Sci 48(1-2): 130-141.

295. Caballero J, Zilocchi S, Tiznad W (2012) Docking and quantitative structure-activity relationship studies for imidazo [1,2-a] pyrazines as inhibitors of checkpoint kinase-1. Med Chem Res 21(8): 1912-1920.

296. Cruzz VL, Martinez S, Ramos J (2014) 3D-QSAR as a Tool for understanding and improving single-site polymerization catalysts. A Review Organometallics 33(12): 2944-2959.

297. Gasteiger J, Marsili M (1980) Iterative partial equalization of orbital electro-negativity: A rapid access to atomic charges. J Tetrahedron 36(22): 3219-3228.

298. Chebib M, Johnston GAR (2000) GABA-activated ligand gated ion channels: medicinal chemistry and molecular biology. J Med Chem 43(8): 1427-1447.

299. Head JD, Zerner MC (1985) A Broyden-Fletcher-Goldfarb-Shanno optimization procedure for molecular geometries. J Chem Phys Lett 122(3): 264-270.

300. Cramer RD, Patterson DE, Bunce JD (1988) Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. J Am Chem Soc 110(18): 5959-5967.

301. Klebe G, Abraham U (1999) Comparative molecular similarity index analysis (CoMSIA) to study hydrogen-bonding properties and to score combinatorial libraries. Comput Aid Mol Des 13(1): 1-10.

302. Wold S, Ruhe A, Wold H (1984) The collinearity problem in linear regression. The partial least squares (PLS) approach to generalized inverses. SIAM J Sci Comput 5(3): 735-743.

303. Aparoy P, Suresh GK, Reddy KK (2011) CoMFA and CoMSIA studies on 5-hydroxyindole-3-carboxylate derivatives as 5-lipoxygenase inhibitors: Generation of homology model and docking studies. Bioorg Med Chem Lett 21(1): 456-462.

304. Singh Khd, Naveena Q, Karthikeyan M (2014) Jak2 inhibitor-a jackpot for pharmaceutical industries: A comprehensive computational method in the discovery of new potent Jak2 inhibitors. Mol Biosyst 10(8): 2146-2159.

305. Yu Z, Li X, Ge C, Si H, Cui L, et al. (2015) 3D-QSAR modeling and molecular docking study on Mer kinase inhibitors of pyridine-substituted pyrimidines. Mol Divers 19(1): 135-147.

306. Golbraikh A, Tropsha A (2002) Beware of Q2. J Mol Graph Model 20(4): 269-276.

307. Wang Z, Cheng L, Kai Z, Wu F, Liu Z, et al. (2014) Molecular modeling studies of atorvastatin analogues as HMGR inhibitors using 3D-QSAR, molecular docking and molecular dynamics simulations. Bioorg Med Chem Lett 24(16): 3869-3876.

308. Mulder GE, Kruijtzer JAW, Liskamp RMJ (2012) A combinatorial approach toward smart libraries of discontinuous epitopes of HIV gp120 on a TAC synthetic scaffold. Chem Commun (camb) 48(80): 10007-10009.

309. Sigrist CJA, Cerutti L, Hulo N, Gattiker A, Falquet L, et al. (2002) PROSITE: a documented database using patterns and profiles as motif descriptors. Brief Bioinform 3(3): 265-274.

310. Sigrist CJA, de Castro E, Cerutti L, Cuche BA, Hulo N, et al. (2013) New and continuing developments at PROSITE. Nucleic Acids Res 41(Database issue): D344-D347.

311. Sabat R, Gruetz G, Warszawska K, Kirsch S, Witte E, et al. (2010) Biology of interleukin-10. Cytokine Growth Factor Rev 21(5): 331-344.

312. Banchereau J, Pascual V, O'Garra A (2012) From IL-2 to IL-37: the expanding spectrum of anti-inflammatory cytokines. Nat Immunol 13(10): 925-931.

313. Gabrysova L, Howes A, Saraiva M, Garra OA (2014) The regulation of apolipoprotein H [Homo sapiens] and pregnancy-associated plasma protein A, pappalysin 1, isoform CRA_b expression. In: Fillatreau S, Garra OA (Eds.), Interleukin-10 in Health and Disease. Curr Top Microbiol Immunol 380: 157-190.

314. Zdanov A, Hihi SC, Wlodawer A (1996) Crystal structure of human interleukin-10 at 1.6 A resolution and a model of a complex with its soluble receptor. Protein Sci 5(10): 1955-1962.

315. Josephson K, Logsdon NJ, Walter MR (2001) Crystal structure of the IL-10 complex reveals a shared receptor binding site. Immunity 15(1): 35-46.

316. DeLano W (2009-2013) The PyMOL Molecular Graphics System, Version 1.6 Schr**ö**dinger, LLC.

317. Wallace AC, Borkakoti N, Thornton JM (1997) TESS: A geometric hashing algorithm for deriving 3D coordinate templates for searching structural databases. Application to enzyme active sites. Protein Sci 6(11): 2308-2323.

318. Barker JA, Thornton JM (2003) An algorithm for constraint-based structural template matching: application to 3D templates with statistical analysis. Bioinformatics 19(13): 1644-1649.

319. Kleywegt GJ (1999) Recognition of spatial motifs in protein structures. J Mol Biol 285(4): 1887-1897.

320. Martin ACR (2000) The ups and downs of protein topology; rapid comparison of protein structure. Protein Eng 13(12): 829-837.

321. Laskowski RA, Watson JD, Thornton JM (2005) Protein function prediction using local 3D templates. J Mol Biol 351(3): 614-626.

322. Yang JM, Tung CH (2006) Protein structure database search and evolutionary classification. Nucleic Acids Res 34(13): 3646-3659.

323. Debret G, Martel A, Cuniasse P (2009) RASMOT-3D PRO: a 3D motif search web server. Nucleic Acids Res 37(Web Server issue): W459-W464.

324. Nadzirin N, Gardiner EJ, Willett P, Artymiuk PJ, Raih MF (2012) SPRITE and ASSAM: web servers for side chain 3D-motif searching in protein structures. Nucleic Acids Res 40(Web Server issue): W380-W386.

325. Nadzirin N, Willett P, Artymiuk PJ, Raih MF (2013) IMAAAGINE: a web server for searching hypothetical 3D amino acid side chain arrangements in the Protein Data Bank. Nucleic Acids Res (Web Server issue): W432-W440.

326. Bahar I, Jernigan RL (1996) Coordination geometry of nonbonded residues in globular proteins. Fold Des 1(5): 357-370.

327. de Araujo AD, Hoang HN, Kok WM, Diness F, Gupta P, et al. (2014) Comparative alpha-helicity of cyclic pentapeptides in water. Angew Chem Int Ed Engl 53(27): 6965-6969.

328. Harrison RS, Shepherd NE, Hoang HN, Gomez GR, Hill TA, et al. (2010) Downsizing human, bacterial, and viral proteins to short water-stable alpha helices that maintain biological potency. Proc Nat Acad Sci USA 107(26): 11686-11691.

329. Goodford PJ (1985) A computational-procedure for determining energetically favorable binding-sites on biologically important macromolecules. J Med Chem 28(7): 849-857.

330. Miller BR III, McGee TD , Swails JM, Homeyer N, Gohlke H, et al. (2012) MMPBSA. py: an efficient program for end-state free energy calculations. J Chem Theory Comput 8(9): 3314-3321.

331. Wang JM, Morin P, Wang W, Kollman PA (2001) Use of MM-PBSA in reproducing the binding free energies to HIV-1 RT of TIBO derivatives and predicting the binding mode to HIV-1 RT of efavirenz by docking and MM-PBSA. J Am Chem Soc 123(22): 5221-5230.

332. Bradshaw RT, Patel BH, Tate EW, Leatherbarrow RJ, Gould IR (2011) Comparing experimental and computational alanine scanning techniques for probing a prototypical protein-protein interaction. Protein Eng Des Sel 24(1-2): 197-207.

333. Greenfield NJ (2006) Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions. Nat Protoc 1(6): 2527-2535.

334. Lakowicz JR (2006) Principles of fluorescence spectroscopy. In: Lakowicz JR (Ed.), Analytical Chemistry. (3rd edn), Springer, New York, USA.

335. Ball C, Vigues S, Gee CK, Poole S, Bristow AF (2001) Rat interleukin-10: production and characterization of biologically active protein in a recombinant bacterial expression system. Eur Cytokine Netw 12(1): 187-193.

336. Künze G, Theisgen S, Huster D (2014) Backbone ^{1}H, ^{15}N and ^{13}C and side chain $^{13}C\beta$ NMR chemical shift assignment of murine interleukin-10. Biomolecular NMR Assignments 8(2): 375-378.

337. Windsor WT, Syto R, Tsarbopoulos A, Zhang RM, Durkin J, et al. (1993) Disulfide Bond Assignments and Secondary Structure-Analysis of Human and Murine Interleukin-10. Biochemistry 32(34): 8807-8815.

338. Kastenholz MA, Pastor M, Cruciani G, Haaksma EEJ, Fox T (2000) GRID/CPCA: A new computational tool to design selective ligands. J Med Chem 43(16): 3033-3044.

339. Papoian GA, Ulander J, Wolynes PG (2003) Role of water mediated interactions in protein-protein recognition landscapes. J Am Chem Soc 125(30): 9170-9178.

340. Teyra J, Pisabarro MT (2007) Characterization of interfacial solvent in protein complexes and contribution of wet spots to the interface description. Proteins 67(4): 1087-1095.

341. Samsonov S, Teyra J, Pisabarro MT (2008) A molecular dynamics approach to study the importance of solvent in protein interactions. Proteins 73(2): 515-525.

342. Klebe G (2015) Applying thermodynamic profiling in lead finding and optimization. Nat Rev Drug Discov 14(2): 95-110.

343. Kutchukian PS, Yang JS, Verdine GL, Shakhnovich EI (2009) All atom model for stabilization of alpha-helical structure in peptides by hydrocarbon staples. J Am Chem Soc 131(13): 4622-4627.

344. Rao T, Ruiz-Gomez G, Hill TA, Hoang HN, Fairlie DP, Mason JM (2013) Truncated and helix-constrained peptides with high affinity and specificity for the cFos coiled-coil of AP-1. PLoS One 8(3): e59415.

345. Case DA, Cheatham TE, Simmerling CL, Wang J, Duke RE, et al. (2012) AMBER 12. University of California, San Francisco, USA.

346. Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, et al. (2006) Comparison of multiple amber force fields and development of improved protein backbone parameters. Proteins 65(3): 712-725.

347. Humphrey W, Dalke A, Schulten K (1996) VMD: Visual molecular dynamics. J Mol Graph 14(1): 33-38.

348. Molecular Operating Environment (MOE), version 2012.10 (2012) Chemical Computing Group Inc. Montreal, QC, Canada.

349. Bayly CI, Cieplak P, Cornell WD, Kollman PA (1993) A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges-the RESP Model. J Phys Chem 97(40): 10269-10280.

350. Dupradeau FY, Pigache A, Zaffran T, Savineau C, Lelong R, et al. (2010) The R.E.D. tools: advances in RESP and ESP charge derivation and force field library building. Phys Chem Chem Phys 12(28): 7821-7839.

351. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, et al. (2009) Gaussian 09, Revision C.01. Gaussian, Inc. Wallingford, CT, USA.

352. (2009) R: a language and environment for statistical computing. R-package Development Core Team.

353. Accelrys Software Inc (2009) Discovery Studio Modeling Environment, Release 25, San Diego, USA.

354. Shepherd NE, Abbenante G, Fairlie DP (2004) Consecutive cyclic pentapeptide modules form short alpha-helices that are very stable to water and denaturants. Angew Chem Int Ed 43(20): 2687-2890

355. (1972) IUPAC-IUB commission on biochemical nomenclature. Symbols for amino-acid derivatives and peptides. Recommendations (1971). J Biol Chem 247(4): 977-983.

356. Kates SA, Albericio F (2000) Solid-phase synthesis: a practical guide. In: Kates SA, Albericio F (Eds.), Marcel Dekker,New York, USA.

357. Amblard M, Fehrentz JA, Martinez J, Subra G (2006) Methods and protocols of modern solid phase peptide synthesis. Mol Biotechnol 33(3): 239-254.

358. Gude M, Ryf J, White PD (2002) An accurate method for the quantitation of Fmoc-derivatized solid phase supports. Lett Pept Sci 9(4-5): 203-206.

359. Kaiser E, Colescott RL, Bossinger CD, Cook PI (1970) Color test for detection of free terminal amino groups in the solid-phase synthesis of peptides. Anal Biochem 34(2): 595-598.

360. Thieriet N, Alsina J, Giralt E, Guibe F, Albericio F (1997) Use of Alloc-amino acids in solid-phase peptide synthesis. Tandem deprotection-coupling reactions using neutral conditions. Tetrahedron Lett 38(41): 7275-7278.

361. http://www.originlab.com

362. Gao X, Han L, Ren Y (2016) In silico exploration of 1,7-diazacarbazole analogs as checkpoint kinase 1 inhibitors by using 3D QSAR, molecular docking study, and molecular dynamics simulations. Molecules 21(5): E591.

363. Dan LL (2012) Tumor heterogeneity and personalized medicine. N Engl J Med 366(10): 956-957.

364. Bucher N, Britten CD (2008) G2 checkpoint abrogation and checkpoint kinase-1 targeting in the treatment of cancer. Br J Cancer 98(3): 523-528.

365. Musk SR, Steel GG (1990) Override of the radiation-induced mitotic block in human tumour Cells by methylxanthines and its relationship to the potentiation of cytotoxicity. Int J Radiat Biol 57(6): 1105-1112.

366. Yan L, Leverson JD (2005) New opportunities in chemo-sensitization and radio-sensitization: Modulating the DNA-damage response. Expert Rev Anticancer 5(2): 333-342.

367. Ma CX, Cai S, Li S, Ryan CE, Guo Z, et al. (2012) Targeting Chk1 in p53-deficient triple-negative breast cancer is therapeutically beneficial in human-in-mouse tumor models. J Clin Investig 122(4): 1541-1552.

368. Wang J, Han X, Zhang Y (2012) Autoregulatory mechanisms of phosphorylation of checkpoint kinase 1. Cancer Res 72(15): 3786-3794.

369. Li T, Christensen SD, Frankel PH, Margolin KA, Agarwala SS, et al. (2010) A phase II study of cell cycle inhibitor UCN-01 in patients with metastatic melanoma: A California cancer consortium trial. Investig New Drugs 30(2): 741-748.

370. Blasina A, Hallin J, Chen E, Arango ME, Kraynov E, et al. (2008) Breaching the DNA damage checkpoint *via* PF-00477736, a novel small-molecule inhibitor of checkpoint kinase 1. Mol Cancer Ther 7(8): 2394-2404.

371. Ma Z, Yao G, Zhou B (2012) The Chk1 inhibitor AZD7762 sensitises p53 mutant breast cancer cells to radiation *in vitro* and *in vivo*. Mol Med Rep 6(4): 897-903.

372. Mitchell JB, Choudhuri R, Fabre K, Sowers AL, Citrin D (2010) *In vitro* **and** *in vivo* **radiation sensitization of human tumor cells by a novel checkpoint kinase inhibitor, AZD7762. Clin Cancer Res 16(7): 2076-2084.**

373. Gazzard L, Appleton B, Chapman K, Chen H, Clark K (2014) Discovery of the 1, 7-diazacarbazole class of inhibitors of checkpoint kinase 1. Bioorg Med Chem Lett 24(24): 5704-5709.

374. Nagro CJD, Jonathan C, Yang X, Linda R, Sankar M, et al. (2014) Chk1 inhibition in p53-deficient cell lines drives rapid chromosome fragmentation followed by caspase-independent cell death. Cell Cycle 13(2): 303-314.

375. Xiao Y, Ramiscal J, Kowanetz K, Del Nagro C, Malek S, et al. (2013) Identification of preferred chemotherapeutics for combining with a chk1 inhibitor. Mol Cancer Ther 12(11): 2285-2295.

376. Gazzard L, Williams K, Chen H, Axford L, Blackwood E, et al. (2015) Mitigation of acetylcholine esterase activity in the 1,7-diazacarbazole series of inhibitors of checkpoint kinase 1. J Med Chem 58(12): 5053-5074.

377. Zhong Z, Liu LJ, Dong ZQ, Lu L, Wang M, et al. (2015) Structure-based discovery of an immunomodulatory inhibitor of TLR1-TLR2 heterodimerization from a natural product-like database. Chem. Commun 51(56): 11178-11181.

378. Ma DL, Chan SH, Wei G, Zhong HJ, Yang H, et al. (2014) Virtual screening and optimization of type II inhibitors of JAK2 from a natural product library. Chem Commun 50(90): 13885-13888.

379. Liu LJ, Leung KH, Chan SH, Wang YT, Ma DL, et al. (2015) Identification of a natural product-like STAT3 dimerization inhibitor by structure-based virtual screening. Cell Death Dis 5: 96-104.

380. Ma D, Chan SH, Leung CH (2013) Drug repositioning by structure-based virtual screening. Chem Soc Rev 42(5): 2130-2141.

381. Ma DL, Chan SH, Leung CH (2011) Molecular docking for virtual screening of natural product databases. Chem Sci 2(9): 1656-1665.

382. Veselinović JB, Nikolić GM, Trutić NV (2015) Monte Carlo QSAR models for predicting organophosphate inhibition of acetylcholinesterase. SAR QSAR Environ Res 26(6): 449-460.

383. Xu C, Ren Y (2015) Molecular modeling studies of [6,6,5] Tricyclic Fused Oxazolidinones as FXa inhibitors using 3D-QSAR, Topomer CoMFA, molecular docking and molecular dynamics simulations. Bioorg Med Chem Lett 25(20): 4522-4528.

384. Cheng LP, Huang XY, Wang Z, Kai ZP, Wu FH (2014) Combined 3D-QSAR, molecular docking, and molecular dynamics study on potent cyclohexene-based influenza neuraminidase inhibitors. Monatshefte Chem 145(7): 9-22.

385. Yanover C, Schueler-Furman O, et al. (2008) Minimizing and learning energy functions for side-chain prediction. J Comput Biol 15(7): 899-911.

386. Sharabi O, Yanover C, et al. (2011) Optimizing energy functions for protein-protein interface design. J Comput Chem 32(1): 23-32.

387. Leaver Fay A, O Meara MJ, et al. (2013) Scientific benchmarks for guiding macromolecular energy function improvement. Methods Enzymol 523: 109-143.