COMPUTATIONAL STUDIES ON INTERACTION BETWEEN SOME RECEPTORS AND LIGANDS OF TRANSFORMING GROWTH FACTOR-BETA SUPERFAMILY:

DESIGN OF INHIBITOR(S) TO PROMOTE OSTEOGENESIS

by

SHAILA AHMED

A dissertation submitted to the Graduate Faculty in Biochemistry in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2010

© 2010

SHAILA AHMED

All Rights Reserved

This manuscript has been read and accepted for the
Graduate Faculty in Biochemistry in satisfaction of the
dissertation requirement for the degree of Doctor of Philosophy.

Date	**Dr. Boojala Vijay B. Reddy** Queens College, **Chair of Examining Committee**
Date	**Dr. Edward J. Kennelly** **Executive Officer**
	Dr. Sreedhara Sangadala (Emory University)
	Dr. Shaneen M. Singh (Brooklyn College)
	Dr. Timothy W. Short (Queens College)
	Dr. Stephane Boissinot (Queens College)

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

COMPUTATIONAL STUDIES ON INTERACTION BETWEEN SOME RECEPTORS AND LIGANDS OF TRANSFORMING GROWTH FACTOR-BETA SUPERFAMILY:

DESIGN OF INHIBITOR(S) TO PROMOTE OSTEOGENESIS

by

Shaila Ahmed

Adviser: Professor Boojala Vijay B. Reddy

Abstract

Transforming growth factor beta (TGF-β) superfamily members execute distinct and intricate roles in numerous biological events such as cell growth, differentiation, embryogenesis, immune responses, and morphogenesis etc. Diverse cellular responses are instigated by binding of these superfamily members to specific transmembrane serine/threonine kinase receptors on the cell surface and thereby activating specific pathways. As a result receptor-regulated Smad proteins are phosphorylated, followed by their complex formation with Co-Smad that together translocates into nucleus and leads to the initiation of transcription of target genes. Bone morphogenic proteins (BMPs), activins, Inhibins are the most important members of TGF-β family. When they exert their biological activity, they are sternly regulated by extracellular antagonists such as noggin, follistatin, CV2, and so forth that are expressed in close temporal and spatial proximity. Blocking these antagonists' interaction with their target receptor proteins helps to promote their respective biological responses when needed. This work is an

attempt to use the current understanding of some of the receptors and ligands of TGF-β superfamily members, namely BMPs, activins and inhibins and their antagonists such as noggin, follistatin, and crossveinless 2 (CV2) interactions through the analysis of their available complex structures to design small molecular inhibitors with an ultimate goal of promoting their respective biological responses.

Noggin is a major natural extracellular antagonist to BMPs which binds to BMPs and blocks binding of them to BMP-specific receptors and thus negatively regulates BMP-induced osteoblastic differentiation. BMPs signal through heteromeric protein complexes composed of type I and type II serine/threonine kinase receptors. Preventing the BMP-2/noggin interaction will preserve free BMP-2 and enhance the efficacy of BMP-2 to induce bone formation. One part of this work is an attempt to use the current understanding of BMP-2, and its interaction with its receptors and antagonists, through the analysis of known structures in protein data bank (PDB), to design inhibitors of BMP-2/noggin interaction with the goal of lowering the dose of BMP-2 required in clinical applications to promote osteogenesis.

Another TGF-β superfamily member, the activin, exerts pivotal roles in male and female reproduction, and has powerful actions in growth and differentiation in various tissues. The ability of activins to assemble their receptor complex, however, is regulated by a number of extracellular binding proteins. Follistatin is one of the main inhibitors among them. Follistatin acts primarily by binding to activin and preventing its interaction with its receptor there by bio-neutralizes activin-mediated responses. Here our main hypothesis is to understand the activin interaction with its receptors and antagonist

follistatin, through the available structures in PDB, to design small molecular binder that blocks activin/follistatin interaction with the goal of promoting BMP responsiveness of the cells.

Crossveinless 2 (CV2), a member of Chordin family, is an extracellular modulator of BMPs which has a unique feature of having both the anti- and pro-BMP activity depending on the cellular context. The anti-BMP activity is directed by excessive dosage of CV2 that impedes the BMP-dependent differentiation of osteoblast and chondrocyte in cell culture and in certain developmental stages. Crystal Structure of BMP-2/CV2-VWC1 domain is available. Here, our main goal is to design a potential inhibitor of BMP-2/CV2 interaction to promote osteoblast differentiation leading to promotion of bone healing. We analyzed the available complex structures to identify contact region of CV2 with BMP-2. By using the binding information of BMP with its receptors from BMP-2/noggin studies, we performed virtual screening and identified several compounds that are likely to bind to CV2 and block BMP-2/CV2 interaction that may enhance the ectopic bone formation.

Finally, by using the BMP-2 sequence fragments that are in contact with noggin, from the BMP/noggin complex structure, we designed several, di-, tri- and tetra-peptide combination structures and docked them onto respective binding sites on noggin with future goal to develop modified peptide mimetic compounds that may help blocking BMP/noggin binding to promote osteogenesis.

To my parents

ACKNOWLEDGEMENTS

From the beginning of this journey one thing became quite clear to me that a researcher can not complete a Ph.D thesis alone. This dissertation is a team effort where many individuals supported me in various aspects. Although the list of individuals I wish to thank extends beyond the limits of this format, I would specially like to thank the following persons for their dedication, prayers, encouragement and support.

First and foremost, I would like to convey my earnest gratitude to my thesis mentor, **Dr. Boojala Vijay B. Reddy**, for his assiduous patience, kind patronage and incessant encouragement. It is a godsend to me to work under his constant attention and guidance throughout this research. Without his effort, it would have been impossible to achieve what is in my hands today. My words fail to thank him enough. I owe him my lifelong indebtedness.

I wish to express my deep appreciation to the members of my thesis committee, **Dr. Sreedhara Sangadala, Dr. Timothy W. Short, Dr. Shaneen M. Singh** and **Dr. Stephane Boissinot**, for their precious time, valuable suggestions and careful counseling to complete the thesis.

It would be incomplete if I don't express my heartfelt thankfulness to all of my colleagues and friends in the lab, especially **Dr. Raghu Prasad Rao Metpally** for his helpful discussions, great inspirations and technical assistance.

I would also like to thank **Dr. Robert Bittman** and **Dr. Wilma Saffran** for giving me the right direction and courage throughout my doctoral studies and for being always there whenever I needed them.

Finally, I am forever indebted to my husband, **Mesbah Ahmed**, my children **Arya** and **Aureeb**, my parents and my siblings for their unconditional love, encouragement and continuous moral support to keep me going and upheld me throughout this journey.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

TGF-β:	Transforming Growth Factor-Beta
BMP:	Bone Morphogenetic Protein
BMPRI:	BMP Receptor Type I
BMPRIA:	BMP Receptor Type IA
BMPRII:	BMP Receptor Type II
ActRIIB:	Activin Receptor IIB
R-Smad:	Regulatory Smad
Smad4:	Common Smad or Co-smad
CV2:	Crossveinless 2
CAP:	Chemicals Available for Purchase
ZINC:	A free database of commercially-available compounds for virtual screening
DS2.0:	Discovery Studio 2.0
LUDI:	*DeNovo* Receptor Docking
Glide:	Grid-based Ligand Docking with Energetics
HTVS:	High Throughput Virtual Screening
SP:	Standard Precision
XP:	Extra Precision
CCDC:	Cambridge Crystallography Data Center
GOLD:	Genetic Optimization for Ligand Docking
%SA:	Percent Solvent Accessibility Contact Area
MM-PBSA:	Molecular Mechanics Poisson-Boltzmann Surface Area

1.0 INTRODUCTION

The TGF-β signaling pathway transduces a variety of extracellular signals into intracellular responses regulating various cellular processes. The disruption of this regulation leads to an array of diseases and developmental disorders [1, 2]. The members of the TGF-β superfamily are structurally related secreted signaling proteins that control various cellular events such as cell-cycle progression, cell differentiation, development, motility, neuronal growth, bone morphogenesis, wound healing, immune surveillance, and so on [1-3]. The vast diversity of ligands of this superfamily gives rise to diverse outcomes depending on each extracellular context [1, 2]. A wide range of cellular responses are initiated by binding of these superfamily members to specific transmembrane serine/threonine kinase receptors on the cell surface and thereby activating them [1, 2, 4]. The stimulated receptors then trigger two types of receptor-regulated Smad (R-Smad) proteins (Smad1/5/8 and Smad2/3) to be phosphorylated [1, 4, 5]. There is a subsequent complex formation of one of the R-Smad types with Co-Smad that translocates into the nucleus and initiates the transcription of specific target genes [1, 2, 4]. In our study, we have focused on ligands, bone morphogenetic proteins, and activins that are the core members of more than 30 TGF-β superfamily growth and differentiation factors [6]. Several extracellular modulator proteins stringently control TGF-β responses among which the major ones are noggin, follistatin, and crossveinless 2 [7-9].

1.1 BMP/Noggin

BMPs are signaling molecules that belong to TGF-β superfamily. More than twenty different BMPs have currently been identified, each of which has its own functions and unique expression patterns depending on amino acid sequence differences [10, 11]. BMPs exhibit roles in the regulation of bone induction, maintenance, and repair. They are important determinants of mammalian embryological development and induce osteoblast differentiation of various types of cells including undifferentiated mesenchymal cells, bone marrow stromal cells, and preosteoblasts [12]. BMP-2 and BMP-4 have been established to be the key factors in embryonic skeletal development [13, 14]. BMP receptors are the transmembrane receptors classified as type I or type II based on sequence homology and they both contain a C-terminal Serine/Threonine protein kinase [15].

BMP ligand binding to type I receptor (BMPRI/ActRI) induces the association of BMPRI and BMPRII receptors, allowing the constitutively phosphorylated BMPRII to phosphorylate and activate the latent BMPRI [16]. After activation of BMPRI, R-Smad is phosphorylated. Phosphorylation of R-Smad releases itself from the receptor complex and forms heterocomplex by associating with Smad4. Subsequently, the R-Smad/Co-Smad complex translocates into the nucleus and regulates the transcription of target genes by functioning in concert with other transcription factors (Fig. 1.1) [16, 17].

BMP activity is tightly regulated prior to receptor recognition by the presence of several structurally distinct extracellular BMP antagonists secreted by the mesenchymal

stem cells, noggin, follistatin, sclerostatin, chordin, DCR, BMPMER, cerberus, gremlin, DAN, and others [16, 18-21]. These are secreted proteins that bind to BMPs and reduce its bioavailability for interactions with the BMP receptors, effectively blocking BMP action. The effects of BMP-2 and BMP-4 are inhibited by the 60-kDa homodimeric protein called noggin [18] known to be expressed in Spemann organizer of the vertebrate gastrula, in neuroectoderm, and later in bone and cartilage. Noggin binds with equal avidity to BMP-2 and BMP-4 and competitively inhibits their interaction with the BMPRIA receptor [19, 20]. The application of BMP-2 or an anti-noggin antibody reversed the effect of exogenous or endogenous noggin, respectively [21] and BMP-2 injection partially cures impaired osteoblast differentiation and bone loss in aged animals [22].

The structural analyses revealed that the surfaces of BMPs have two prominent hydrophobic patches, the convex type II receptor-binding interface (knuckle epitope) and the concave type I receptor-binding interface (wrist epitope) [23]. The residues determining binding affinity and specificity of BMP-2 and BMPRIA were identified by mutational analysis. These studies suggested that the elusive binding determinants of the wrist epitope for the BMP-2/BMPRIA interaction involved the mainchain amide groups of amino acids L51 and D53 with a minor contribution from the hydrophobic interactions. Two hydrogen bonds are formed between the L51 residue (main chain amide and carbonyl) of BMP-2 and the Q86 residue of BMPRIA [15]. The core structure of BMPRII shares the same pattern of disulfide connectivity as ActRIIB [10]. In an earlier report a hypothetical complex between BMPRII and BMP-2, created by superimposing BMPRII in the exact position of ActRIIB within the ternary complex

between ActRIIB, BMP-2, and BMPRIA, suggested that BMPRII makes very similar overall ligand binding contacts to BMP-2 as does the ActRIIB receptor. The residues Y67, W85, and F115 of BMPRII are within the main hydrophobic patch while the residues K81, S86, E93, and Y113 are also participate as the important binding determinants. In addition, the residues H87 and Y40 confer specificity in BMPRII ligand binding [10]. Previous mutational analyses identified A34, H39, S88, L90, and L100 residues as binding determinants of BMP-2 for both the BMPRII and ActRIIB receptors [15].

The superposition of the noggin/BMP-7 complex structure onto a model of the BMP signaling complex shows that noggin binding effectively masks both the pairs of binding epitopes. Noggin contains an extended N-terminal 'clip' segment consisting of about 20 residues [15]. The type I receptor-binding site is obstructed by a segment of the clip domain (Q28 to D39). The hydrophobic ring of P35 of noggin inserts into a hydrophobic pocket on BMP-7 formed by the residues, W52, W55, V87, Y128, and M131, that mimics a similar insertion of F85 from BMPRIA into a hydrophobic cleft on BMP-2. The type II receptor-binding site is masked extensively by the C-terminal half of the clip segment (N40 to E48) [15, 23]. Another mutational study revealed that noggin mutants, with substitutions at positions L46D, E48K, and I218E, essentially abolished the BMP-7 binding activity in some of the variants and the P35R substitution diminished the affinity for BMP-7 [7].

The orthopaedic surgical treatment of musculoskeletal problems frequently requires bone grafting to promote healing. Spinal fusion surgery, the joining (or fusing)

of one or more vertebrates to reduce pain and stabilize the spine, is a common treatment for spinal disorders such as low back pain, degenerative spinal diseases, and non-healing bone fractures. The FDA approved BMP-2 (rhBMP-2, InFUSE of Medtronic) for its use in anterior inter-body bone fusion and BMP-7 (OP-1 of Stryker Biotech) to treat non-unions in long bones where auto graft is unfeasible and alternative treatments have failed. However, cost prohibitive high dose requirements due to their antagonists limit their use in patients. The large dose of rhBMP-2 may also lead to anterior cervical spine complications. To overcome this limitation, it is necessary to utilize additional strategies that can potentiate BMP activity and thus enable a reduction of BMP doses required for clinical use. In this study the first part of work was undertaken with an ultimate goal of functionally blocking the physiological antagonist noggin towards potentiating the BMP activity.

Together with the structural knowledge about the binding determinants of BMPs, BMP receptors, and noggin found from previous mutational studies, we initially analyzed and explored the available structures of BMPs, its receptors, and noggin to identify the various residues involved in their interaction. In section 3.1, we present our results on the analyses of all these receptor-ligand and ligand-antagonist complex structures and on the identification of binding hotspots on the noggin surface. We further used the LUDI *de novo* design method, Glide, and the GOLD docking methods to virtually screen a large number of compounds against the binding sites of noggin to identify the lead chemical compounds that block specifically only the BMP-2/noggin interaction but not the BMP-2 interaction with its receptors. Since the BMP-2/4 binding regions of BMP receptors and the BMP-2/4 binding regions of noggin are very similar, we identified the small

molecules that bind to noggin binding epitopes but do not bind to the BMP receptor epitopes. We further plan to test the binding affinity and the biological activity of these molecules using biochemical and cell-based assays in future. Here in section 3.2 we present our computational analysis and molecular docking studies to select potential molecules that bind noggin from the available drug like small molecular databases. We carried out only the modeling and the virtual screening studies and section 3.2 shows our results in these analyses.

1.2 Activin/Follistatin

Activin, a member of the TGF-β superfamily, is involved in growth and differentiation of many tissue types and is crucial for both male and female reproduction by regulating the release of follicle-stimulating hormone (FSH) from the pituitary [24-26]. Activin-A is a homodimer consisting of two disulfide-linked polypeptide chains, termed β_A, of 116 amino acids in length [27]. Activin initiates signaling through two types of transmembrane serine/threonine receptor kinases classified as type II (ACVR2A, ACVR2B or also abbreviated as ActRIIA, ActRIIB) and type I [or activin like ligands (ALK); ACVR1 (ALK5), ACVR1B (ALK4) and ACVR1C (ALK7)] [28]. Upon binding of activin-A to the type II receptor [29], type I receptor is recruited into the complex and is phosphorylated by type II receptor [30, 31]. Activated type I receptor then phosphorylates Smad proteins. The activated Smads translocate to the nucleus and activate transcription of the target genes (Fig. 3.1) [8, 24, 32].

Activin, secreted in a biologically active form as 25 kDa protein, is regulated by a number of high and low affinity extracellular binding proteins that determine binding to

its receptors [33, 34]. The chief among the activin binding proteins is follistatin, a multi-domain protein [35] that binds activin with sub-nanomolar affinity and renders it to be incapable of interacting with its receptors [36] thereby bio-neutralizing the activin mediated responses. Follistatin knockout mice show defects in size, skeletal, and muscle development, display improper tooth and whisker development. The animal skin is taut and shiny, they fail to breath and die shortly after birth [37-39]. Follistatin, other than potent inhibition of activin, also binds to other TGF-β ligands, such as inhibin, BMPs 2, 4, 6, 7, 11, and 15, and myostatin with lower affinities that is consistent with the above phenotypic effects resulting from follistatin regulating multiple TGF-β ligands *in vivo* [8, 37, 38, 40, 41].

Follistatin is a glycoprotein, expressed as two splice variants (isoforms), one ending at residue 288 (FS288) and the other ending at residue 315 (FS315) [42]. These two variants differ in their C-terminal sequences and are generated by a combination of alternative splicing and proteolytic processing. The FS315 is a major circulatory form [34, 43], whereas the FS288 form associates with the cell surface and is involved in activin internalization and degradation [44]. Both forms of follistatin bind activin-A with high affinity and thereby lead to a diminished bio-availability of activin [24, 45-47]. The mature follistatin is composed of an N-terminal unique domain, FSD0 and three follistatin domains (FSD1, FSD2, FSD3) shown in Fig. 3.2 (B). Each follistatin domain is further divided into an epidermal growth factor-like (EGF-like) and a Kazal protease inhibitor-like sub-domains [34, 48, 49]. The main functional difference between the isoforms is in their ability to bind heparin sulphates (HS). The shortest form FS288 has the highest affinity for HS and the longest form FS315 with its acidic tail fails to bind HS

[50, 51]. The FSD1 domain bears a highly basic region in its EGF-like sub-domain which is the main binding site for HS [52, 53].

The activin/follistatin complex consisting of two activin molecules and a follistatin molecule (Fig. 3.2 (B)), was crystallized in a dimer form consisting of four activin and two follistatin molecules [54]. Each of the two FS288 molecules extends around an activin subunit, forming head-to-tail contacts between their N- and C-terminal domains [8]. The FS N-terminal domain contacts the wrist region of one activin-A subunit and the concave fingers of the second activin-A subunit [15] while FSD1, FSD2, and FSD3 contact only one activin subunit, extending around its outer knuckles and fingertips [8] (Fig. 3.2 (B)). In the crystal structure of the ActRIIB extracellular domain bound to activin-A, the receptor makes contact with the convex outer face of the finger region of the ligand. The activin-A/ActRIIB interface involves the hydrophobic and the ionic/polar residues in Finger 1 and Finger 2 of activin-A (Fig. 3.2(A)). The residues in Finger 1 of activin-A are F17, I30, A31, P32, and H36 and in Finger 2 are R87, P88, S90, L92, Y94, I100, K102, and E111 [55]. Even though there is no crystal structure available for activin-A bound to ALK4, the activin residues essential for binding to the type I receptor were identified using a selective mutagenesis approach and these were the residues, M91, I105, and M108 on the concave surface of Finger 2 and the S60 and I63 residues in the α-helix [47, 56, 57]. The activin/follistatin crystal structure indicated that the follistatin contacts two contiguous surfaces on activin-A [8, 47] and one of the surfaces overlaps the type II receptor binding site. The residues I30, A31, P32, L92, Y94, I100, and K102 are the residues on the convex outer side of the activin β-strands corresponding to the high-affinity binding site identified by mutagenesis and the residues

D95-N99 are in the β-strand finger tips. The second surface maps to the proposed type I receptor binding site. This is comprised of the most of the wrist region of activin-A including the residues from the concave β-strand of one subunit (W25, W28, M91, Y93, and I105) and the helical region of the second subunit (H47, I48, G50, T51, S52, F58, T61, and H65). The prediction was that the follistatin would inhibit both type I and II receptor binding by burying these residues on activin [8, 47].

To validate the importance of the type I and type II interfaces for follistatin binding, activin-A mutants were generated by mutating residues to alanine at these sites and on surrounding surfaces. The mutants were then evaluated for their ability to bind follistatin and ActRII and checked how the mutants behaved in terms of inducing FSH secretion by the anterior pituitary. The residues that were mutated at the activin-A / ActRIIB interface, or knuckle region which is the convex outer face of the finger region of activin-A showed conspicuous effects on follistatin binding. The high-affinity follistatin binding determinants of activin are the residues Y94 and I100 to a greater extent, and the residues L92 and K102 to a lesser extent of activin. The mutation of the I30 and A31 residues effected protein expression. The mutation of S90 is disruptive for ActRIIB binding but did not impact follistatin binding which suggesting that the interaction surfaces for type II receptors and follistatin are overlapping but not identical [47].

The mutagenesis of the residues to alanine within the β-strand fingertips (D95-E99), the helical regions (P45-I48; S52-S56; S60-H65; and M68-H71) and the concave surface of the fingers (M91, Y93, and M108) of activin-A exhibited the following results.

The affinity lowered noticeably for the mutation of D95, but the mutation of the surrounding residues had no effect on follistatin binding. This indicates that follistatin wraps around the fingertips of activin-A but the direct contacts in this region are minimal [47]. The mutational approach at the wrist region of activin-A identifies the residues as being not necessary for a high-affinity follistatin binding with a very little effect on follistatin binding upon mutations in the α-helical region. The mutation of residues M91 and Y93 to alanine in the concave surface of fingers disrupts the juxtaposed binding surface used by ActRII and follistatin on the convex surface of finger 2, consequently lowering activin's affinity for both ActRII and follistatin. Another mutation in the concave surface, M108A, also exhibits no effect on follistatin binding either. The analyses of all these mutagenesis studies indicate that the residues at the convex outer surface of the fingers of activin-A are relatively the most important high-affinity binding determinants for follistatin and as compared to the wrist region of the activin-A [47].

From the crystal structure, it is known that the N-terminal domain of follistatin, which is required for biological activity, mimics the BMP type I receptor and occupies the proposed type I binding surface of activin-A. From various functional analyses it is known that the follistatin constructs FSD1-FSD2 and FSD1-FSD2-FSD3 exhibit almost identical affinity for activin, especially the main epitope for activin binding resides within the FSD2 domain. The interaction surfaces for FSD1-FSD2 and type II receptor overlap nearly perfectly. The FSD2 domain interacts with a large hydrophobic surface in the knuckle epitope and with the activin fingertips. The R192 is an important residue within the FSD2 domain participating in the interaction between FSD1-FSD2 and activin-A. The hydroxyl group of Y94 from the activin knuckle region is hydrogen-bonded to the

mainchain oxygen of R192 of the FSD2 domain and the side chain of R192 is bonded with the polar residues D27 and Q98 of activin fingers. The R192A mutation abolishes this interaction whereas the wild-type FSD1-FSD2 domain showed a significant binding affinity (a k_d of 430 nm) in a previous isothermal titration calorimetry experiment [51].

The inhibition of follistatin binding to activin helps in promoting the TGF-β and BMP signaling process and modulates the intricacy of several differentiation and growth related disease conditions. In this study, we initiated the analyses of various available structures of the activin, its receptors, and follistatin to identify the key residues involved in their interaction from structural details and the mutational studies. Next, we used the docking procedure such as LUDI *de novo* design method, Glide, and GOLD docking to computationally screen a large number of compounds against the binding sites of follistatin to obtain the high ranking chemical compounds that potentially block only the activin-follistatin interaction but not the activin interaction with its receptors. In this study, we portrayed the virtual screening studies and molecular docking experiments to search for the possible blockers of follistatin from the readily available drug like small molecular databases. We intend to determine the binding affinity and the biological activity of these molecules using several biochemical and cell-based assays. The section 3.3 gives results from our structural and docking studies.

1.3 BMP-2/Crossveinless-2 (CV2)

The TGF-β superfamily member BMP, mediates numerous biological processes from dorso-ventral patterning to organogenesis and cellular differentiation in the

developing embryo and the post embryonic life in vertebrates and invertebrates through extracellular signaling [58-62]. In section 1.1, we described BMPs that exert their biological activity by binding to transmembrane receptor complexes of type I and type II Ser/Thr kinases, thereby turning on the BMP responsive genes in the nucleus in concert with the other transcriptional factors. There are a series of activation and phosphorylation events of the kinase receptors and Smad proteins prior to entering the nucleus [58, 61, 63, 64] (Fig. 1.1). In section 1.1, we discussed the structural features of BMP-2 and BMPRs in detail about the convex type II receptor-binding interface (knuckle epitope) and the concave type I receptor-binding interface (wrist epitope) [23], and their binding determinants [10, 15].

The Crossveinless-2 protein (CV2) (Fig.4.1) is secreted at high BMP signaling sites and controls the activity of BMP during the dorso-ventral patterning of *Drosophila* and *Xenopus* embryos, during the formation of *Drosophila* wing and mouse organogenesis [61-63, 65-67]. The CV2 can potentiate or antagonize the BMP generated activity gradients by transmitting signals that determines the differentiation and fate of recipient cells by inducing multiple threshold responses [61, 62, 68]. The extracellular modulator, CV2, is essential for the survival at birth to assist and establish morphogenetic fields of BMP signaling [62]. The CV2 is a member of the Chordin family. In contrast with other modulators such as noggin or follistatin, the Chordin family has a unique feature of performing two opposite functions in interacting with the members of TGF-β superfamily. This family include both members that exhibit anti- and pro-BMP activity depending on the cellular context [61, 63-66]. Lowering of the CV2 levels can convert CV2 from an antagonist to an agonist of signaling making this protein a biphasic one.

Different levels and types of BMPs can influence this biphasic activity [69]. Previous mRNA injection and mutagenetic studies provide evidence that the anti-BMP activity of over expressed CV2 is confined to the CV2-VWC1 domain *in vivo* whereas the pro-BMP activity of CV2 is very likely to be independent of binding to BMP where only the VWC1 domain is inadequate [9]. The excessive doses of CV2 inhibited the BMP-dependent differentiation of endothelial cells, osteoblast, and chondrocyte in cell culture *in vitro* and in certain developmental stages [64, 65] *in vivo*. Another anti-BMP role of CV2 is observed in *Xenopus* where CV2 operates as a local feedback inhibitor [67]. However, *in vivo* experiments showed a positive BMP/CV2 feedback loop during the wing development in *Drosophila* [70], zebrafish gastrulation [63] and in mouse organogenesis [61]. The proteolytic cleavage converts CV2 as a pro-BMP factor from an anti-BMP factor. The ability of zebrafish CV2 to enhance BMP signaling requires the cleavage of CV2 whereas *Drosophila* CV2 does not need the proteolytic cleavage to carry out the same function [63, 69]. The CV2, which was first identified as a BMP modulator protein in *Drosophila*, contains five VWC domains at its N-terminal followed by a C-terminal segment comprising a single VWD and a trypsin inhibitor-like cysteine-rich (TIL) domain (Fig. 4.1). The CV2 protein binds with BMP with 2:1 stoichiometry through its first N-terminal VWC1 domain that contains a rod-like structure with an overall dimension of 50Å×25Å whose amino acid sequence is highly similar in Drosophila, Xenopus, zebrafish, chicken, and man [9]. The remaining four other VWC domains 2–5 and the VWD and TIL domains do not interact with BMP-2 [9, 71].

Both modulator proteins bind with BMPs via VWC domains, occurring several times [21, 71, 72] with high affinity and play an important roles in development and

diseases [61]. The truncation studies show that only the first of five VWC domains of CV2 determines the binding activity to BMP-2 [71]. An investigation of the individual VWC domains of different proteins reveals that the VWC domain (also called Chordin-like cysteine-rich or CR domain) is a versatile binding module that can render an array of binding specificities in its multiple forms and environments [71] in the BMP and other TGF-β signaling pathways as extracellular modulators [73]. The VWC domain is found in about 500 extracellular proteins from *Drosophila* to human which typically contains less than 100 residues and contains a conserved C*XX*C*X*C and/or CC*XX*C sequences in common while the rest of the sequences are highly diverse among them [74].

The CV2-VWC1 domain is a tripartite structure comprised of an N terminal clip (L1–S8), an N-terminal subdomain SD1 (C9–K42), and a C-terminal subdomain SD2 (C43–G66) [9] (Fig. 4.2). The N-terminal clip is disulfide bonded to the SD1 subdomain showing an extended conformation with its C-terminus. The SD1 domain exhibits a three-stranded all-antiparallel β-sheet named β1, β2, and β3 and the SD1 fold are stabilized by two disulfide bonds: one between β1 and β2 and the other between β2 and β3. The SD2 domain contains a short two-stranded antiparallel β sheet named β4 and β5 that lacks a secondary structure. However, the SD2 domain is stabilized by two disulfide bonds that crosslink the long β3 β4 loop to the N- and C-terminus of β5 making the SD2 structure to be rigid. Finally, there is one more disulfide bond that ties the SD1 and SD2 domains making the whole structure even more compact [9].

The previous structure analysis unveils that the CV2-VWC1 domain inhibits types I and II receptor binding concomitantly. The N-terminal clip itself prevents BMP-2 from

binding with the type I receptor by hindering most of the hydrophobic patch of the wrist epitope of BMP-2 needed for the type I receptor binding. The residues L1–E6 from N-terminal clip fit snugly into the cleft comprised of helix 1 of one BMP-2 monomer and the β-sheet of fingers 1 and 2 of the other BMP-2 monomer. These interactions mostly involve mainchain atoms hydrogen bond network between clip and BMP-2, thereby stabilizing the complex due to a highly cooperative binding [9, 71]. In comparison with the N-terminal clip, the SD1 domain appears to be responsible for the most of the BMP-2 inhibition as it occupies the majority of the type II receptor binding site and thereby obstructs type II receptor binding to BMP-2. It is known from the CV2-VWC1/BMP-2 structure that the SD2 domain, consisting of 24 residues, is conserved among BMP binding VWC domains which lacks the direct contact with BMP-2 molecule and might modulate BMP-activity in some yet undefined way [9].

The CV2-VWC1/BMP-2 complex structure portrays that despite being tiny, the clip and SD1 assures high-affinity binding between the proteins via strong cooperativity when they are linked together. The contribution of the clip to BMP-2 binding was evaluated by constructing two CV2-VWC1 truncated variants which lack the residues L1–T5 or L1–A7. These two variants showed a 100- and 2000-fold lower affinity for BMP-2, respectively, which proves high-affinity binding to BMP-2. The truncated variant of SD2 of CV2-VWC1 binds BMP-2 with wild-type-like affinity demonstrated no input to binding of BMP-2. SD1 dealings with the knuckle epitope of BMP-2 provide only low-affinity binding similar to the type II receptor-BMP-2 interaction whereas the clip region shows low nanomolar affinity [9]. The results from the truncation and mutagenesis studies exhibit significant diminution in the binding affinity of N-terminal

clip of CV2-VWC1 for BMP-2 [9]. The variant I2A shows 6-fold lower affinity for BMP-2 binding, validating the inevitability of the hydrophobic knob-into-hole interaction between I2 of VWC1 to BMP-2. The variant T5P displays a 20-fold lower affinity for BMP-2 authenticating the necessity of the H bond formation between the main chain groups of T5 of VWC1 with N102 of BMP-2. Another important mainchain H bond interaction is between A7 of VWC1 and L100 of BMP-2 which contributes a 20-fold difference in binding affinities for BMP-2 upon two truncations in the CV2-VWC1 clip region that were mentioned above. The variant T3P shows only a 2.5-fold lower affinity to BMP-2 showing only a marginal contribution. The other interactions between the CV2-VWC1 N-terminal clip and BMP-2 appear to be energetically silent [9].

Some mutagenesis studies were performed to verify the contribution of the VWC1 SD1 domain interactions with the knuckle epitope of BMP-2 showing an extensive reduction in binding affinity of SD1 domain for BMP-2 upon mutation of the residues I21 or I27 to Ala and I21 and A36 to Arg. About 1.3 kcal mol^{-1} of binding free energy is lost owing to I21A or I27A mutation whereas a very tight packing is observed when there were even more severe mutations such as I21R and A36R [9].

In section 1.1, we summarized the BMPs' role in orthopedic surgical treatments. However, to reduce the cost and to avoid side effects, large dose of BMP is not recommended. Our strategy is to overcome these constraints by blocking the physiological antagonist, CV2 protein that would potentiate the BMP activity and augment the use of BMP for clinical use.

In concert with the structural information about the binding determinants of BMP-2, BMP receptors, and CV2-VWC1 domain attained from preceding mutational studies, we began analyzing and exploring the available structures of BMP-2, its receptors and the CV2-VWC1 domain to detect the various residues that are in direct contact. Then a number docking procedure such as LUDI *de novo* design method, Glide and GOLD docking were brought into play to virtually screen a large number of compounds against the binding sites of the CV2-VWC1 domain to acquire lead chemical compounds that potentially block only the BMP-2/CV2 interaction but not the BMP-2/BMPRs interactions. Thus, we carried out our computational study and molecular docking experiments to select credible molecules that bind CV2 from the accessible drug like small molecular databases. In this document we provided results from modeling and virtual screening studies in section 3.4. We intend to ensure the binding affinity of these molecules using various biochemical and cell-based assays currently under development.

1.4 Peptido-mimetic Compounds

Peptido-mimetic compounds are small designer drug-like compounds and are alluring alternative to small molecules provided by various available databases [75, 76]. These mimetic compounds with their unique distribution of amino acid residues are expected to show specific binding to the target protein [77]. The peptides that are constructed by simply joining the amino acid residues one-by-one have to go through modifications to alter its stability or biological activity to be effective as therapeutic agents. These modifications to the peptide will not occur naturally; we have to improve the secondary and tertiary structures of these peptides so that it can be docked onto the

receptor protein competently and efficiently [75]. Hence, these modified peptides can provide fabulous avenues to design drug-like molecules that may mimic the structurally and conformationally important epitopes involved in protein-protein interactions [78].

In our study, we used BMP-2 sequence fragments derived from the key binding residues that are in contact with noggin, from the BMP-2/noggin complex structure and modeled several, di-, tri- and tetra-peptide of variable combinations. Then we used molecular docking techniques to dock these peptides onto respective binding sites on noggin and evaluated the scores. Our future plan is to suggest modified peptide mimetic compounds that may be used as therapeutic drugs in blocking BMP/noggin binding to promote osteogenesis and to lower the expenses in bone related treatments.

2.0 MATERIALS AND METHODS

Our objective is to analyze TGF-β ligands/receptors and TGF-β ligand antagonists and their interactions from the known structures of these complexes and use this knowledge to design small molecules that block the binding of antagonists to TGF-β ligands. The employed methods include computational analysis of TGF-β ligands, their receptors' and antagonist's sequences, structures, and various docking and computational screening procedures. For BMP/noggin, we analyzed BMP-2/noggin, BMP-2/BMPRs to design small molecules that block the antagonist noggin binding to BMP-2 by various computational analyses and docking procedures using BMP-2, BMPRs and noggin sequences and structures. We used the human sequences from SWISS-PROT [79] database. The sequences and structures are: BMP-2 (P12643), BMP-7 (P18075), noggin (Q13253), BMPRIA (P36894) and BMPRII (Q13873). The structures from Protein Data Bank (PDB) - BMP-2 (1rew), BMP-7 (1m4u), noggin (1m4u), BMPRIA (2h62), ActRIIB (2h62) and BMPRII (2hlr) - were used.

For activin/follistatin, we explored activin/follistatin and activin/ActRIIB extracellular domain interactions from known crystal structures of these complexes. Since there is no crystal structure available for activin-A bound to ALK4, the employed methods include only the computational analysis of activin, ActRIIB, and follistatin structures, and various virtual screening procedures. The structures of activin (1s4y), follistatin (2b0u), and ActRIIB (1s4y) were used from PDB.

Finally, for BMP-2/CV2, we only had to analyze BMP-2/CV2-VWC1 interactions from known structures of the crystal complex (PDB ID: 3bk3) to design

small molecules that block the antagonist CV2 binding to BMP-2. We are using the same analysis and docking experiments for BMP-2/BMPRs interaction that were performed for BMP/noggin experiments. We used the ZINC [80] and the CAP (Accelrys Inc.) small molecular databases of commercially available compounds. The 2,667,437 molecules classified as purchasable in the ZINC-7 database were used from the website: http://zinc.docking.org/index.shtml. We also used 70,000 molecules in the CAP database supplied by Accelrys Inc. and carried out our virtual screening experiments, separately on each system using each database.

2.1 Amino Acid Sequence Alignment

The sequences of BMPs, BMP receptors and noggin were retrieved from their structures. A multiple sequence alignment for these sequences was generated using ClustalW [81] to identify equivalent residues (Fig. 1.2) that were used for modeling BMP-2/noggin complex based on available BMP-7/noggin complex structure.

2.2 Solvent Accessible Contact Area Calculations

TGF-β ligands, TGF-β receptors, various antagonist structures and their complexes were analyzed to identify the key residues involved in their interaction by solvent accessible contact area calculation using the method of Richmond & Richards (1978) [82]. The method gives the surface area ($Å^2$) for each residue that comes into contact with a spherical probe of radius of 1.4 Å. The percentage difference in solvent accessible contact area of each residue in a protein on contact with another protein of the complex pair gives information about how much area of a residue is involved in binding.

For BMP/noggin, the analyses were performed on BMP-2, BMP Receptors, and the antagonist, noggin structures; for activin/follistatin, the analyses were on activin, ActRIIB and the antagonist follistatin structures; for BMP-2/CV2, only the BMP-2/CV2-VWC1 structure complex was analyzed.

2.3 General Utility Software

We have written a few utility programs in the process of analyses: (i) A program which clusters the high scoring 500-1,000 molecules based on their distribution on the surface of the antagonist binding site such as the surface of noggin, follistatin and CV2 binding site to facilitate selecting the molecules targeted to different hotspot locations on the surfaces. (ii) A program to compute consensus scores with appropriate conversion of Glide-XP score and GoldScore of each molecule to a common scale. The program normalizes the Glide-XP score to the same scale as GoldScore to arrive at a CombScore given in every table of Glide and GOLD docking results as shown in section 3.

2.4 Graphical Viewing and Superposing the Protein Structures

We used Accelrys Inc. Discovery Studios 2.0 (DS2) tools for various computational needs such as: to display, to superpose, and to generate molecular graphics figures. The sophisticated application modules in Discovery Studio, Schrodinger and Cambridge Crystallography Data Center's (CCDC) software enabled us to dock and score ligands and to design lead compounds and analyze the results. Schrodinger's Maestro module provides an intuitive, full-featured building tool for constructing molecular models of any type. This allows constructing new molecules from molecular

fragments or individual atoms. It also allows modifying such properties as atom type, hybridization, potential function parameters, bond order, and geometry of existing molecules.

2.5 LUDI *De Novo* Receptor Docking

LUDI is a method for the *de novo* design of ligands for proteins [83, 84]. The LUDI score is a sum of five contributions: from ideal hydrogen bonds, contributions from perturbed ionic interactions (interaction of donor/acceptor in the receptor, e.g. COO-, or NH3+), contributions from lipophilic interactions, contributions due to the freezing of internal degrees of freedom of the ligand, contributions due to the loss of translational and rotational entropy of the ligand. In receptor mode, LUDI searches a fragment library for complementary small molecules that best bind the defined interaction site(s). During the search and fit computation, LUDI also determines the energy estimates, or scores, for each conformation searched for the fragments in the library. In our study we used LUDI receptor mode against TGF-β ligands, their receptors, and their antagonists which screened potential lead compounds from the CAP and the ZINC small molecular weight compound databases. LUDI fits small molecules and fragments into the receptor sites by matching complementary polar, hydrogen bonding and hydrophobic groups. LUDI is not so powerful in active docking and scoring ligands but extremely useful in suggesting modifications to the known ligands by scoring candidate derivatives in the receptor binding site that may enhance their binding. It offers significant time savings in the search for the new and potentially improved ligands and is also useful in identifying hotspot binding sites within the protein-protein binding surfaces.

2.6 Glide Docking

Schrodinger's Glide offers a full spectrum of speed and accuracy from high-throughput virtual screening of millions of compounds to very accurate binding mode predictions, providing consistently high enrichment at every level [85-87]. Maestro is Schrödinger's powerful, unified, multi-platform graphical user interface. Protein preparation is usually required for Glide calculations which can be performed for the most protein and protein-ligand complex structures from PDB. The quality of the results obtained from Glide depends critically on the quality of the starting structures. Receptor Grid Generation is performed to set up and start a grid file calculation function. The Grid files represent physical properties of a volume of the receptor, specifically the active site that will be searched when attempting to dock a ligand. The shape and properties of the receptor are represented on a grid by several different sets of fields that provide progressively more accurate scoring of the ligand poses. In Ligand Docking, Glide provides a rational workflow for virtual ligand database screening from HTVS (high throughput virtual screening) to SP (standard precision) to XP (extra precision). The XP offers the most accurately docked poses and the highest level of enrichment and requires considerably more CPU time [86]. To handle a large set of ligands we screened them in SP, for rapid screening we use HTVS, and then the high scoring structures from those runs are docked by XP for final refinement.

2.7 GOLD docking

GOLD is a powerful genetic algorithm for docking flexible ligands into protein binding sites. Predicting how a small molecule will bind to a protein is difficult, and no

method can guarantee the success. The next best option is to measure as accurately as possible the reliability of the method, i.e., the chance that it will make a successful prediction in a given instance. For that reason, GOLD has been tested on a large number of complexes extracted from the Protein Data Bank. The overall conclusion of these tests was that the high ranking GOLD solution was correct in 70-80% of cases [88-90]. GOLD provides all the functionality required for docking ligands into protein binding sites from prepared input files. GOLD will likely be used in conjunction with a modeling program to create and edit starting models. The commonly used molecular modeling environments include: SYBYL and Insight II or Cerius2. GOLD docks each ligand several times starting each time from a different random population of ligand orientations. The results of the different docking runs are ranked by fitness score. The number of dockings to be performed on each ligand is set when the ligand file is defined and by default the number of dockings to be performed on each ligand is 10. The total time spent in docking a ligand obviously depends on the number of docking runs, so you can make GOLD run faster by reducing this number. The GOLD offers a choice of scoring functions, such as GoldScore, ChemScore, Astex Statistical Potential (ASP), and User Defined Score. With respect to using the GoldScore, ChemScore or ASP scoring functions one may provide a successful prediction where the other fails, but their overall success rates are about the same. Different values of the genetic algorithm parameters may be used to control the balance between the speed of GOLD and the reliability of its predictions. The GOLD fitness function is made up of four components as follows: protein-ligand hydrogen bond energy (external H-bond), protein-ligand van der Waals (vdw) energy (external vdw),

ligand internal vdw energy (internal vdw), and ligand torsional strain energy (internal torsion)[88-90].

2.8 Insight II and Discovery Studios

Insight II (Accelrys Inc.) is a sophisticated molecular modeling environment initially developed on Silicon Graphics systems to provide a powerful graphical interface with the best-of-breed algorithms for molecular dynamics, homology modeling, *de novo* design, and electrostatics—making it the useful solution for protein modelers, computational chemists, and structural biologists. Most of the modules of Insight II are made available in Discovery Studios (Accelrys Inc.). Several key programs are already available, including CHARMm, MODELER, Profiles-3D, and LUDI (2001-2007 Accelrys Software Inc). Small molecules can be built and/or modified using Builder or Biopolymer modules. The molecule is constructed by bonding together the fragments that collectively constitute a single entity. The fragments, stored in Insight II libraries, represent a wide variety of functional groups. The Biopolymer module facilitates the building and modification of peptides, proteins, polynucleic acids, and carbohydrates. In particular, for peptides there are various options such as build the peptide sequences while imposing secondary structure, delete or replace residues in peptides and proteins, impose secondary structure on existing peptides and proteins, and change N and C terminal capping groups etc.

2.9 Impact

Impact is a molecular mechanics and dynamics program that provides the molecular mechanics component of Glide, Liaison, and QSite calculations. Some of the basic capabilities of Impact are also available from Maestro. These capabilities include molecular mechanics (MM) energy minimization, molecular dynamics (MD) simulations, hybrid Monte Carlo (HMC) simulations, and addition of explicit water solvent to a structure. Maestro panels for Basic Impact applications can be opened from the Impact submenu of the Applications menu. The Impact Energy Minimization panel is used to set up and run an Impact energy minimization calculation on the structure in the Workspace. The Energy Minimization panel has three tabs: Potential, Constraints, and Minimization. The Potential tab sets parameters that control how Impact calculates the molecular-mechanics energy in a minimization calculation or dynamics simulation. The Constraints tab of the Impact Energy Minimization panel is used to set up Impact *atom constraints.* These include constraints that freeze selected atoms at their input coordinates (frozen atoms) or keep them near their initial coordinates by applying a harmonic force (constrained atoms). The basic settings of the Impact energy minimization task are defined in the Minimization tab [91]. The Impact was used for energy minimization of di-, tri- and tetra peptides generated for peptide mimetic docking studies.

2.10. MM_PBSA Method to Study Ligand-Protein Interaction

The Molecular Mechanics_Poisson-Boltzmann Surface Area (MM-PBSA) method was used to study ligand-protein interaction of the ranked molecules from the Ludi, Glide and Gold docking. We have used the Discovery Studio 2.0 CHARMm-based methods for estimating the translational, rotational and vibration entropies of protein-

ligand systems. The MM-PBSA and MM-GBSA are emerging as useful and effective approaches. The results showed that correlations between MM-PBSA or MM-GBSA binding free energies with experimental affinities were in most cases excellent agreement, Importantly the correlations obtained with the use of a single protein-ligand minimized structure and with implicit solvation models were similar to those obtained after averaging over multiple MD snapshots with explicit water molecules, with consequent save of computing time without loss of accuracy. When applied to a virtual screening experiment, such an approach proved to discriminate between true binders and decoy molecules and yielded significantly better enrichment curves [96]. In our virtual screening we have also computed the MM-PBSA energies for the molecules selected and ranked.

3.0 RESULTS

3.1 Modeling and Analysis of BMP and its Complex Structures with Receptors and Antagonist, Noggin.

Since BMP-2/noggin complex structure was not available, the crystal structure of BMP-7 and noggin complex was used to model BMP-2/noggin complex. Similarly, from the crystal structure of ternary complex of BMP-2/BMPRIA/ActRIIB (2h62) ectodomain [23] a new ternary complex BMP-2/BMPRIA/BMPRII was modeled. In this section, the modeling and analysis of these complexes is described to identify various binding sites on the BMP-2, BMPRs, and noggin. The noggin binding surface using the LUDI *de novo* docking procedure with the CAP database of small molecules to identify the hotspot regions where small molecules may find docking sites.

3.1.1 Identification of BMP-2/Noggin Binding Regions

The equivalent residues of BMP-2 that interact with noggin and *vice versa* were identified from the crystal structure of BMP-7 and noggin complex (1m4u). The BMP-7, BMP-2, and noggin are homologous structures. We therefore used ClustalW [81] to identify the corresponding equivalent residues among these sequences (Fig. 1.2). Though the BMP and noggin exhibit similar structures and the major regions of their sequences are similar, their multiple sequence alignment showed that the BMP binding region to noggin is not equivalent to the noggin binding region to BMP. The BMP-7 protein sequence is highly homologous to BMP-2 with 80% of sequence identity. The crystal structure of BMP-2 dimer was therefore superimposed onto the BMP-7 dimer of the

complex consisting of BMP-7 and noggin by using a web based tool Superpose [92] . We have then modeled BMP-2 dimer structure in the position of BMP-7 dimer and energy minimized the BMP-2/noggin model complex (Fig. 1.3(B)) through Impact energy minimization using default parameters. From the calculation of percent solvent accessibility contact area (%SA) of this new model complex in the presence or absence of the binding partners we identified the interacting residues of noggin with BMP-2 and the *vice versa* (Column 1 and 4 of Table I.1).

3.1.2 Identification BMP-2/BMPRIA/BMPRII Binding Regions

From the crystal structure of ternary complex of BMP-2/BMPRIA/ActRIIB ectodomain [23] a new ternary complex BMP-2/BMPRIA/BMPRII was modeled. The core structure of BMPRII ectodomain (2hlr) is that of the three-finger toxin fold shared by other TGF-β family receptors. The BMPRII receptor shares the same pattern of disulfide connectivity as ActRIIB, with disulfide bonds between pairs of cysteines in positions 34–66, 60–84, 94–117, 99–116, and 118–123 (BMPRII numbering) with very similar structure. The overall structure has a core of antiparallel β-sheet formed by the three longest β-strands. A complex between BMPRII and BMP-2 was modeled by superimposing BMPRII in the position of ActRIIB of the ternary complex of BMP-2/BMPRIA/ActRIIB. The energy of the newly modeled ternary complex BMP-2/BMPRIA/BMPRII was minimized through Impact energy minimization using default parameters (Fig 1.3(A)). The percent solvent accessible contact area of each residue was calculated in monomer structures of the model ternary complex of BMP-2/BMPRIA/BMPRII to identify the interacting residues. In Table I.1 the amino acid

residues of both the receptors BMPRIA and BMPRII that bind to BMP-2 and the amino acids of BMP-2 that bind both these receptors are given.

3.1.3 Comparison of BMP-2/BMPRs and BMP-2/Noggin Interactions

The superposition of BMP-2 proteins in the BMP-2/noggin complex and BMP-2/BMPRIA/BMPRII model complex structures shows that noggin binding effectively masks both pairs of binding epitopes of BMP-2 for its type I and type II receptors (Fig 1.3). The type I receptor-binding site is occluded by the residues of a segment of the clip domain (M27, Y30 to D39), which is bound predominantly through hydrophobic interactions by the residues from Q208 to R210. The hydrophobic ring of P35 of noggin inserts into a hydrophobic pocket of BMP-2 mimicking a similar insertion of F85 from BMPRIA into a hydrophobic cleft of BMP-2. This residue is thought to be a key determinant of all type I receptors. The type II receptor-binding site is masked extensively by the residues from the C-terminal half of the clip segment (N40, L43, L46, I47, H49), the residues from finger 2 (H199, R204, R206, Q208 to R210, I218 to Q221, P223) and the residues from the N-terminal clip segment (P37 to D39). Thus, the two motifs by which noggin binds BMP-2 have marked resemblances to those of the receptors: first, a hydrophobic side chain from a flexible backbone segment is inserted into the hydrophobic pocket of BMP; and second, a set of complementary interactions occur between the two curved hydrophobic surfaces. An additional resemblance is that noggin, as a homodimer, binds to both pairs of receptor-binding epitopes of BMP-2, mimicking closely the binding mode of the hetero-tetrameric receptor assembly.

A comparison of the BMP-2/noggin complex with the BMP-2/BMPRIA structure shows that noggin covers a part of the wrist epitope with a short stretch of ten residues from the N-terminal clip segment (Q28 to D39). The binding determinants of BMP-2 for noggin and the type I receptor are quite different. The main epitope of BMPs for binding to noggin will overlap with the 'knuckle' epitope. Similar fractional accessibilities in the noggin and ActRIIB contact exist for many BMP-7 residues. A closer inspection shows that noggin binds more towards the outer edge of finger 2 as compared to the binding site of the BMPRII receptor. In Table I.1 we have given the percentage solvent accessibility of each residue of noggin, BMPRIA and BMPRII involved in binding with BMP-2. In Table I.2 we have given the BMP-2 residues involved in interaction with the residues (number) of noggin, BMPRIA, and BMPRII and these are the key residues finally utilized to select small molecules.

3.1.4 Identification of Potential Small Molecular Binding Hotspots on the Noggin

The solvent accessible contact area of the noggin surface covered by BMP-2 binding was calculated to be about 484 $Å^2$. It can be seen from Fig. 1.4 that the noggin binding surface has several concave grooves where small molecules can find tight binding hotspots. The noggin surface region was further explored to identify binding hotspots using LUDI by employing the default docking parameters; some advanced parameters we used are given in Table I.3. The LUDI analyses found discrete positions in the clefts of protein structures suitable to form hydrogen bonds or to fill a hydrophobic pocket. We targeted the CAP small molecular databases against the 484 $Å^2$ areas of the noggin regions that bind to BMP-2. The docking studies described below show that there

are at least three hotspot locations around the residues at position L46, R204 and I220 on the structure of noggin surface where most of the small molecules cluster in space to bind onto the noggin surface (Fig. 1.6).

3.2 Virtual Screening and Selection of Drug-like Compounds that Block Noggin Interaction with BMPs to Promote Osteogenesis.

LUDI *de novo* design method of Accelrys Discovery Studios 2.0 was used against the noggin surface regions that bind to BMP-2. The BMP-2 binding region of noggin was divided into four sub groups (with some overlapping residues) for LUDI runs (Fig. 1.5). Group I consisted of nine amino acids M27, Y30, L31, H32, I33, R34, P35, A36, P37; group II consisted of seven amino acids, PSDN from 37-40, L43, L46, and P223; group III consisted of eight amino acids I47, H49, H199, R204 and IPIQ from position 218 to 221; group IV consisted of four amino acids R206 and QRR at position 208 to 210. The binding site in each of these four groups was defined as a sphere of radius each of 10.5, 10.5, 12, and 9 Å for group I, II, III and IV, respectively (Fig. 1.5). Each group covered between 85 $Å^2$ to 200 $Å^2$ binding surface area.

3.2.1 Screening Noggin Binding Surface Using the CAP Database and the LUDI Docking

The LUDI search was carried out against the four regions described above using the 70,000 molecules of the CAP database which yielded about 11,088 molecules with Ludi3 scores greater than threshold score of 300. The LUDI search against BMPRIA and BMPRII binding region for BMP-2 yielded 10,074 and 10,643 molecules, respectively.

We also performed LUDI screening against the region of BMP-2 that binds to both the receptors and obtained a set of 15,031 molecules. The small molecules that are common to either of the receptors (BMPRIA or BMPRII), BMP-2, and noggin were removed so that the small molecules that do not bind to receptors and BMP-2 but only binds to the inhibitor noggin were used for further screening. The common molecules found from BMP-2, BMP receptors and noggin runs were excluded from the 11,088 molecules and the remaining 2,283 molecules were selected that were unique to noggin binding. Using only these 2,283 small molecules, we generated a small library of compounds and using this two LUDI runs were performed on the entire BMP-2 binding region of noggin with more stringent LUDI search parameters. As a result, we obtained 1,306 small molecules with Ludi3 binding scores that were greater than the 300 threshold. To rank and select these molecules, we used Ludi3 scores and the solvent accessible contact areas of ligands with binding residues to identify molecules that were distributed over the entire binding region of noggin surface (Table II.1). To our surprise, we found that the majority of these molecules were clustered in concentric locations. A set of 100 high ranking molecules with significantly high Ludi3 scores are shown in Fig. 1.6. Binding poses for some of the high scoring molecules are shown in Fig. 2.1.

3.2.2 Screening Noggin Binding Surface using the ZINC Database and the LUDI Docking

We followed the same procedure to screen the ZINC database against the noggin regions that bind to BMP-2 and identified a total of 51,222 small molecules. LUDI analyses against the residues of BMPRIA and BMPRII that bind to BMP-2 generated

50,181 and 33,724 molecules respectively with score greater than 300. We also conducted the LUDI search against the residues of BMP-2 that bind to both the receptors and found 22,147 molecules. A set of 4216 common molecules, found from BMP-2, BMP receptors and noggin runs, were excluded from 51,222 small molecules and the remaining 47,006 molecules were noted as unique to noggin. These molecules were sorted and 11,587 molecules were chosen that have Ludi3 scores 250 or higher. By using Prepare Ligands module of DS 2.0 a library of 10,758 non-redundant ligands was prepared. We then carried out one LUDI run on the entire BMP-2 binding region of noggin using radius of 16 Å sphere. As a result, we found 4,059 small molecules with Ludi3 binding scores that were greater than 300. We identified molecules with high scores from the ZINC database that were distributed over the entire binding region. A very small number of selected high scoring molecules covering the entire binding region are given in Table II.2. The docking poses of a few of these selected molecules are shown in Fig. 2.2.

3.2.3 Screening BMP-2, BMPRIA and BMPRII and Noggin Binding Surface Using the ZINC Database and the Glide Docking

Since noggin is a large molecule, its entire binding region is divided into two sub-regions, and two separate HTVS runs were performed. Single HTVS runs were performed on the selected residues of BMP-2, BMPRIA, and BMPRII. To define the active site for receptor grid generations we selected the centroid of the residues in the binding regions of noggin (two regions), BMP-2, BMPRIA and BMPRII in the corresponding runs. Both the CAP and ZINC databases were used with default Glide

parameters by setting the number of atoms as 200 and rotatable bonds to 30 as constraints to select the ligands. Finally, in the output, we limited the total number of the predicted best-binding poses per docking run. For HTVS, we kept the default option as 10,000 and for SP and XP only the best 1,000 high scoring molecules were considered for ranking.

For noggin we performed Glide search in HTVS mode by selecting all binding residues and obtained a set of 10,000 high scoring molecules. For the selected residues of BMP-2, BMPRIA, and BMPRII, we performed single HTVS and obtained the best scoring 10,000 molecules for each run. From these HTVS runs using the ZINC database, the unique small molecules found against noggin, BMP-2, BMPRIA, and BMPRII were about 9248, 9995, 10,000, and 9,999, respectively. We sorted the 27,879 unique molecules among BMP-2, BMPRIA and BMPRII, and 1,991 common molecules among noggin, BMP-2, BMPRIA and BMPRII. These 1,991 small molecules were removed from the total unique noggin blockers so that the small molecules that do not bind to BMP-2 and both of its receptors but only binds to the inhibitor noggin were used for further Glide runs. Thus, we found 7,257 exclusive noggin blockers from the ZINC database and with these we performed SP and selected the best scoring 1,000 molecules from SP output to use for XP run to get the final XP scoring.

3.2.4 Screening BMP-2, BMPRIA and BMPRII *and Noggin Binding Surface Using the CAP Database and the Glide Docking*

We followed the same procedure for the CAP database as mentioned above. The CAP molecules obtained against noggin, BMP-2, BMPRIA, and BMPRII were 10,000 for each protein. The unique molecules among BMP-2, BMPRIA, and BMPRII were

22,220 and the common molecules among noggin, BMP-2, BMPRIA, and BMPRII were 5715. After excluding these 5715 molecules from unique noggin binding molecules, we found 4284 exclusive noggin blockers from the CAP database and performed SP and XP mode docking as explained previously. These docking procedures generated a set of report files and pose viewer files along with other files where ligands were arranged so that poses with the highest scores were sorted to higher ranks in the list. These highest scored ligands were expected to show the most favorable interaction with the receptor proteins.

3.2.5 Screening BMP-2, BMPRIA and BMPRII *and Noggin Binding Surfaces Using the ZINC and CAP Databases and the GOLD Docking*

Screening of the ZINC and the CAP databases was performed using GOLD. Since GOLD takes considerably large amount of time, we used the 1,000 top scoring molecules from the Glide XP runs by making a ligand library for GOLD docking. PDB input files of single chains of BMP-2, BMPRIA, BMPRII and noggin were used. Then the binding site was defined from a list of all binding residues for each separate protein in text files. To add a ligand we had to format the ligand file acceptable for the GOLD run. GOLD docked each ligand for 10 times and then the results of different dockings were ranked by the fitness score. We chose the GoldScore as a fitness function and kept all other parameters as default values. We obtained a summary of the fitness scores for all the docking attempts on a specific ligand which was in the decreasing order of the fitness score and selected the molecules that show the best scores.

3.2.6 Consensus Scores from the Glide and the GOLD Docking

We computed a consensus score using the Glide and GOLD scores together. Glide gives a fitness score as negative integer value, the smaller the value the better is its score. The Gold Score is a positive integer and the higher the value the better is its score. Among the best scoring molecules, we finally obtained a range between the highest to the lowest scoring molecule in both the scores separately and gave a 50% weight to each score to normalize to 100 as maximum possible value as consensus combined score (CombScore). We then calculated equivalent CombScore for each molecule using respective Glide and GOLD scores. An excel file was created with Glide, GOLD, CombScore values and the solvent accessible contact area of each ligand molecule to the noggin residues in the binding region. Using all these parameters we selected a set of molecules with high Glide, GOLD, and CombScore that were distributed over the hotspot binding region and some of these molecules are shown in Table II.3 and II.4. In future experimental studies these molecules could be test for their binding to noggin by biochemical assays. In Fig. 2.3 and 2.4 binding poses of some of the selected high scoring molecules are shown. The molecules VCGN001 to VCGN006 with CombScores of 123, 122, 118, 117, 115 and 115, respectively are shown in Fig. 2.3. They exhibit a hydrophilic region forming two or more hydrogen bonds and molecule VCGN005 hydrophobic region is pocketed into the hydrophobic grove giving highest CombScores. In Fig. 2.4 the molecules VZGN004 to VZGN005 with CombScores of 126 and 125 show 3 and 5 hydrogen bonds and molecules VZGN001 to VZGN003 show strong hydrophobic and hydrophilic aromatic groups binding at their respective complementary location, in different orientations showing both the hydrophobic and hydrophilic regions pocketed into the respective complementary grooves with at least one hydrogen bond. In

Fig. 2.5, it was shown that the amino acids, L46, E48 and I218 and P35, whose mutations fail the noggin binding to BMP-2 [16]. In our docking studies we identified a significant number of molecules binding to these residues that may result in blocking the noggin binding to BMP-2. Finally we have calculated the binding energy of all the selected molecules using MM-PBSA approach and the resulting values are given in Table II.5.

3.3 Structural Analysis and Virtual Screening for Potential Inhibitor(s) of Activin/Follistatin Interaction to Promote Activin Mediated TGF-β and BMP Signaling.

Crystal structures of activin/follistatin (Fig. 3.2(B)) and activin-A/ActRIIB extracellular domain (ED) (Fig. 3.2(A)), available from human, were used in this analysis to analyze various binding epitopes. Here we present our studies that lead to selection of a manageable number of high scoring follistatin binding small molecular blockers which could be used in future for experimental validation of their binding through biochemical assays.

3.3.1 Identification of Follistatin and Activin Receptor Binding Regions

From the crystal structures of complexes the amino acid residues of ActRIIB-ED and follistatin binding to activin were identified by calculating surface accessibility of ActRIIB-ED and Follistatin in the presence or absence of respective interacting partner as described in the method. Similarly, the amino acid residues of activin that binds to ActRIIB-ED were identified. Since activin/activin receptor type I (ALK-4) structure is not available in PDB, we used experimentally available residue information from the

literature. From the activin/follistatin complex structure it was calculated that a total of 508 $Å^2$ region of follistatin binds to activin. In this binding region of activin approximately 177 $Å^2$ was covered by the same residues of activin that bind to ActRIIB-ED. The surface accessible contact area in $Å^2$ and percentage of total residue contact area involved in binding of each residue are shown in Table III.1. In Fig. 3.3(ADF), we have shown the entire binding surface region of follistatin with activin. It can be seen from Fig. 3.3 that primarily the ND, FSD1, FSD2, and partially FSD3 domains are involved in binding with activin. As can be seen from the Fig. 3.3 (BCE), the binding surface of follistatin contains a few charged atomic regions having possibility for hydrogen bonding with small molecules. We also see quite a few hydrophobic concave scaffolds where hydrophobic surface of small molecules may be used as binding pockets.

In Table III.2, the residue number of follistatin and ActRIIB binding to activin are shown, and these are the residue regions that were finally used to target small molecules. The Table III.2 shows that the residues, I30, A31, P32, and H36 from finger 1 and the residues, R87, P88, S90, L92, I100, and K102 from finger 2 of activin interact with ActRIIB. Table III.2 also shows that the residues, I30, A31, P32, L92, I100, and K102 from activin β-strand and D95-N99 from activin β-strand fingertips interact with follistatin which overlaps with the type II receptor binding site. The residues W28, Y93 and I105 from one of the activin β-strand subunit and the residues, H47, I48, G50, T51, S52, F58, and T61 from the helical region of the second activin subunit interact with follistatin which map to the proposed type I receptor binding site. From Table III.2 we also see that follistatin residues, L5, R6, Q14, V15, L16, Y17, N44, L46, F47, K48, M50, I51, and R192 bind with the above residues of activin that mimics the type I binding site.

On the other hand the follistatin residues, L105, D106, E123, P125, E126, V151, V152, D153, Q154, Y159, V161, L191, R192, T195, C196, S201, G203, and L204 mimics the type II receptor binding site of activin.

In order to apply virtual screening the entire activin binding region of follistatin was divided into four groups with some overlapping residues as shown in Fig. 3.4. The Group I consisted of amino acids at position R6, Q14, V15, L16, Y17, N44, L46, F47, K48, M50, I51, and F52; group II consisted of amino acids at position L105, D106, E123, P125, E126, and Y159; group III consisted of amino acids at position L5, R6 and R31; group IV consisted of amino acids at position V151, V152, D153, Q154, V161, L191, R192, T195, C196, S201, G203, and L204. The binding sites were defined as spheres of radius of 14, 13, 8, and 14 Å covering the amino acid in the Group I, II, III, and IV, respectively.

3.3.2 Screening Follistatin Binding Regions Using the ZINC Database and the LUDI Docking

LUDI *de novo* design method of Accelrys Discovery Studios 2.0 was used against the follistatin regions that bind to activin. The LUDI search against the four regions with the ZINC database gave 12,085 molecules with Ludi3 scores greater than 300. The follistatin binds to a significant number of same residues of activin that binds either type I receptor or ActRIIB. The LUDI was used against the ActRIIB surface that binds to activin and obtained a set of 9,804 unique ActRIIB potential binding molecules that show significant binding score of 300. We also used the LUDI against activin residues that bind to both the receptors, the ActRIIB-ED and ALK4, and obtained 8,897 potential

activin binding molecules. The ALK4 binding region obtained from literature, consists of the residues, M91, I105, and M108 on the concave surface of finger 2 and the residues, S60 and I63 in the α-helix [47, 56, 57]. We obtained 14,552 unique blockers of activin and ActRIIB, and 4,832 common molecules among follistatin, activin and the ActRIIB. The common molecules found from activin, ActRIIB-ED and follistatin runs were excluded from these molecules and the remaining 7,253 molecules were considered specific to follistatin binding from which we removed redundant ligands and found 7,220 molecules that were unique to follistatin binding. Using these 7,220 small molecules, we generated a mini library and used for three LUDI runs on the entire activin binding region of follistatin. From these runs 4,275 small molecules were obtained, with many common molecules, with Ludi3 binding scores greater than 300. Similar to the BMP/noggin studies, we selected a set of high scoring molecules from the ZINC database that were distributed over the entire binding region of follistatin and a small number of high ranking molecules are shown in Table III.3, the binding poses for few selected molecules are shown in Fig. 3.6.

3.3.3 Screening Follistatin Binding Regions Using the CAP Database and the LUDI Docking

We pursued the same procedure as mentioned before to screen the CAP database and identified a total of 4,465 small molecules against follistatin binding surface. We obtained 4,056 molecules against ActRIIB that bind to activin and 3,701 molecules against activin residues that bind to both the receptors. These molecules were sorted and 5,766 unique binders of activin and ActRIIB were obtained with 1,990 common

molecules between them. The LUDI search against these four regions (Fig. 3.4) of follistatin yielded 4,465 molecules with Ludi3 scores greater than 300. The common molecules found from activin, activin receptor and follistatin runs were excluded and the remaining 2,444 molecules, specific to follistatin binding, after removing redundant molecules, were selected to generate a mini library consisting of these molecules. Using these molecules, three LUDI runs were performed further on the entire activin binding region of follistatin using spheres of different radius by employing the advanced LUDI parameters. These runs gave 1,758 small molecules, many common among different runs, with Ludi3 binding scores greater than 300. We created an excel file as mentioned before with a combination of Ludi3 scores and the solvent accessible contact areas of ligands with binding residues. A set of molecules with high scores that were distributed over the entire binding region are shown in Table III.4. The binding poses of a limited number of these molecules are shown in Fig. 3.7

In Fig. 3.5, we displayed about 150 high scoring molecules from the LUDI screenings from the entire region but most of the high scoring molecules were concentrated in the three clusters showing that there are at least three hotspot regions on follistatin to target the small molecules that can effectively bind to follistatin. In Fig. 3.5, we have shown the distribution of the best scoring 50 molecules around each of the three binding hotspots. In the final selection of molecules we gave high weightage to each of these three binding hotspot regions. In the Table III.3, III.4, III.5, and III.6, we listed the molecules pertaining to each of these three hotspot regions from the LUDI, Glide combined with Gold scoring, respectively.

3.3.4 Screening Follistatin Binding Regions Using the ZINC Database and the Glide Docking

Follistatin is a big molecule; therefore, with the experience of LUDI screening, the entire binding region of follistatin was divided into three regions shown in Fig. 3.5. Three separate Glide runs were performed on these three regions. However, single runs were performed on the activin and ActRIIB binding regions. In Glide receptor grid generations, we selected the centroid of the residues of each binding region of follistatin (three regions), ActRIIB, and activin in the corresponding runs to define the active sites. Like the BMP/Noggin studies, both the CAP and the ZINC databases were used along with mostly default Glide parameters and constraints. In the output tab we limited the total number of the predicted best binding poses per docking run, for HTVS as 10,000 high scoring molecules and for SP and XP we set 1,000 high scoring molecules. For follistatin, we performed three separate Glide HTVS runs by selecting all the residues in three hotspot regions and obtained a set of 10,000 molecules from each run. We performed single HTVS runs for the selected residues of ActRIIB and activin, and obtained high scoring 10,000 molecules for each run. From these HTVS runs, the unique small molecules found against all three separate follistatin runs, ActRIIB and activin were 10,000 each. Then we ran Glide-SP and selected the highest scoring 500 molecules for each follistatin run and 1,000 for ActRIIB and activin to use for XP run to get the best scoring molecules. Then we sorted and obtained 108, 166, and 133 common molecules among the regions (i), (ii) and (iii) of follistatin, activin, and ActRIIB, respectively. These small molecules were removed from the unique 500 follistatin blockers from each group so that the small molecules that do not bind to activin and its receptor but only binds to

the inhibitor follistatin were selected as exclusive follistatin blockers. We found 392, 334 and 367 exclusive follistatin blockers for region (i), (ii) and (iii), respectively, from the ZINC database. A list of a few high ranked molecules was generated and some are given in Table III.5 and the docking poses for few of these molecules are shown in Fig. 3.8.

3.3.5 Screening Follistatin Binding Region Using the CAP Database and the Glide Docking

The CAP database was screened against follistatin binding region using the same methods as stated before. The unique CAP molecules found against all three separate follistatin runs, ActRIIB and activin were 10,000 for each protein. Then we performed Glide-SP and Glide-XP as described before and obtained 500 molecules for each follistatin run and 1,000 molecules each for ActRIIB and activin run. A set of 231, 226 and 203 common molecules among the group I, group II, and group III of follistatin, activin and ActRIIB, respectively, were removed from the unique 500 follistatin blockers from each group. After excluding the common compounds, we found 56, 74, and 83 exclusive follistatin blockers for group I, group II and group III, respectively. Some selected high ranking molecules from these runs are given in Table III.6 and the poses of a few of them are given in Fig. 3.9.

3.3.6 Screening Follistatin Binding Region Using the ZINC and the CAP Databases and the GOLD Docking

Further screening of the ZINC and the CAP databases was performed with the GOLD docking. We used high scoring 500 molecules for follistatin and 1,000 each for

ActRIIB and activin from the Glide XP runs and used them as ligand libraries for GOLD docking. The PDB input files of single chains of follistatin, ActRIIB, and activin were used here. The same parameters were used as before for BMP/noggin studies and we performed the GOLD docking run in background mode. The final set of selected high scoring molecules from the ZINC and the CAP database are given Table III.5 and Table III.6, respectively, and the binding poses of a few of these molecules are shown in Fig. 3.8 and 3.9, respectively.

3.3.7 Consensus Scores from the Glide and the GOLD Docking

Similar to the strategy employed for BMP-2/noggin blockers, we computed a consensus CombScore using the Glide and the GOLD scores together. A set of molecules with high Glide, GOLD, and CombScore scores were selected that were distributed over the entire binding region (Table III.5 and III.6). In Fig. 3.8 and 3.9 we displayed the binding poses for a few of these high scoring selected molecules. From previous site directed mutation studies it was observed that R192A mutant of follistatin failed to bind activin [51]. From Tables III.3, III.4, III.5 and III.6 we see that this residue is located on hotspot III region which is potentially a very good target site for small molecules. We therefore identified a few high scoring molecules in the region surrounding this key residue that may block this residue (shown in some poses in Fig. 3.6, 3.7, 3.8 and Fig. 3.9). The molecules given in tables, and also additional molecules, upto 30 under each screening, can be tested for their binding with follistatin in biochemical and subsequent cell-based assays. Finally we have calculated the binding energy of all the selected molecules using MM-PBSA approach and the resulting values are given in Table III.7.

3.4 Structural Analysis and Virtual Screening for Potential Inhibitor(s) of BMP-2/CV2 Interaction to Promote Osteogenesis

Analysis of CV2-VWC1/BMP-2 complex structure was carried out to identify binding regions on the CV2-VWC1 to design small molecules that block the key amino acid residues on CV2 to prevent it from binding to the BMP. The crystal structure of the ternary complex between BMP-2/BMPRIA/BMPRII was also analyzed to identify various binding regions. In this study several high scoring CV2 blockers were identified which could be used in future for experimental confirmation.

3.4.1. Identification of BMP-2/CV2-VWC1 and BMP-2/BMPRs Binding Regions

Using the crystal structure of BMP-2/CV2-VWC1 (3bk3) domain we calculated percent solvent accessibility contact area (%SA) of this complex in the presence and absence of the binding partners and identified the interacting residues of CV2-VWC1 with BMP-2 and *vice versa*, and these are given in Column 1 and 3 of Table IV.1. The interacting residues of BMP-2 that communicate with both the receptors and *vice versa* were analyzed in the BMP/noggin studies (Table I.1).

3.4.2. Comparison of BMP-2/BMPRs and BMP-2/CV2 Interactions

The crystal structure of the BMP-2/CV2-VWC1 complex and the BMP-2/BMPRIA/BMPRII model complex structures shows that VWC1 domain binding simultaneously inhibits BMP-2 from binding with its type I and type II receptors (Fig. 4.2 and Fig. 1.3(A)). The N-terminal clip of CV2 blocks the type I receptor binding site of

BMP-2 which fits into the cleft formed by the helix 1 of one of the BMP-2 monomer and β sheet of fingers 1 and 2 of the other BMP-2 monomer (Fig. 1.3(A) and Fig. 4.2). The clip residue I2 represents another binding site by pointing into a cleft which is similar to the knob-into-hole interaction in BMP-2/BMPRIA, where the residue, F85 of BMPRIA points into the hydrophobic pocket created by the residues, W28, M89, Y103, M106, and I62 in BMP-2. The type II receptor binding site is masked extensively by the residues from the SD1 subdomain. The hydrophobic interactions mostly from hydrophobic residues dominate this compact binding epitope. The residues, A7, I18, I21, and A36 of CV2-VWC1 totally bury the residue, L100 of BMP-2. V33, A34, P35, V98, V99, and L90 from BMP-2 were other crucial hydrophobic residues for type II receptor binding providing evidences that the interaction in the knuckle epitope of BMP-2 depends solely on hydrophobic forces (Table IV.2).

A comparison of the BMP-2/VWC1 complex with the BMP-2/BMPRIA structure shows that VWC1 covers most of the hydrophobic patch of the wrist epitope with a short stretch of six residues from the N-terminal clip segment (L1 to E6). The relatively flat VWC1 binds to BMP-2 like paperclip holding a paper sheet, with subdomains SD1 and SD2 sticking onto the knuckle epitope and the clip extending from SD1 folding over into the wrist epitope of BMP-2 (Fig. 4.2). The binding determinants of BMP-2 for VWC1 and the type I receptor overlaps with the wrist epitope whereas the main epitope of BMPs for binding to VWC1 will overlap with the knuckle epitope of BMP-2 which provide the bulk of the interaction for BMP-2 inhibition. In Table IV.1 the percentage solvent accessibility of each residue of CV2-VWC1, BMPRIA and BMPRII involved in binding with BMP-2 are given. In Table IV.2 the BMP-2 residues involved in interaction with the

VWC1, BMPRIA and BMPRII amino acids are given, and these were the residue regions finally considered for selecting the target small molecules.

3.4.3. Identification of Potential Binding Hotspots on the CV2-VWC1 Domain

The solvent accessible contact area of the CV2-VWC1 surface covered by BMP-2 binding was computed to be about 271 $Å^2$. Fig. 4.3 shows that the CV2-VWC1 binding surface contains quite a few concave grooves where small inhibitor molecules can find tight binding spaces. We examined the CV2-VWC1 surface region to identify binding hotspots using the LUDI and the Glide docking. These methods located distinct positions in the clefts of the receptor CV2-VWC1 domain suitable for the small molecules to form hydrogen bonds or to fit in hydrophobic pockets. Our objective was to screen the CAP and the ZINC small molecule databases against the 271 $Å^2$ binding surface of the CV2-VWC1. The docking studies with the LUDI *de novo* design method and the Glide using the CAP and the ZINC small molecular database library (described below) showed that there was a hotspot location around the residues T5, A7, I21, I27 and A36 on the structure of the CV2-VWC1 surface where most of the high scoring molecules positioning themselves (Fig. 4.5).

3.4.4. Screening CV2-VWC1 Surface Using the CAP and the ZINC Databases and the LUDI Docking

The LUDI docking was carried against the CV2-VWC1 surface regions that bind to BMP-2. The BMP-2 binding region of CV2 is divided into two sub groups for LUDI runs (Fig. 4.4). Group I consisted of eight amino acids W-1, L1, I2, T3, G4, T5, E6 and

A7; group II consisted of seven amino acid I18, N20, I21, I27, A36, C38 and Q40. The binding site in these two regions was defined as spheres of radius 12.0 Å and 10.0 Å for group I and II, respectively (Fig. 4.4). Each region is covered between 100 $Å^2$ to 170 $Å^2$ binding surface area.

The LUDI screening against these two regions using the CAP database yielded 1,906 molecules with the Ludi3 scores greater than the 300 threshold. From earlier studies, the LUDI search against the BMPRIA and BMPRII binding regions to the BMP-2 had yielded 10,074 and 10,643 molecules, respectively and against the region of BMP-2 that bind to both the receptors, had yielded 15,031 molecules. The small molecules that were common to both the receptors (BMPRIA or BMPRII), the BMP-2 and the CV2-VWC1 were removed so that they do not bind to the receptors and BMP-2 but bind only to the CV2-VWC1 and were used for further screening. The redundant 1,817 common molecules found from BMP-2, BMP receptors and CV2-VWC1 analyses were excluded from the 1,906 molecules of CV2-VWC1 and the remaining 89 molecules were unique to the CV2-VWC1 binding. A set of 100 high scoring molecules with high Ludi3 scores are shown in Fig. 4.5. Using only these 89 small molecules we have generated a mini library and one LUDI search was performed on the entire BMP-2 binding region of CV2-VWC1 using 15 Å-radius of sphere employing the more stringent LUDI search parameters. We obtained 36 small molecules with Ludi3 binding scores with greater than the 300 threshold. We computed solvent accessible contact area to identify molecules with high scores that were distributed over the entire binding region of the CV2-VWC1 surface (Table IV.3) and docking poses of some of them are shown in Fig. 4.6.

The ZINC database was screened against the CV2-VWC1 regions that bind to BMP-2 in the same manner and identified a total of 4,606 small molecules. From earlier analyses described in section 1.2, the LUDI searches against the binding surfaces of the BMPRIA and BMPRII to the BMP-2 generated 50,181 and 33,724 molecules, respectively; against the residues of BMP-2 that bind to both the receptors generated 22,147 molecules with the Ludi3 scores greater than 300. A set of 3,627 common molecules found from BMP-2, BMP receptors and CV2-VWC1 searches were excluded from the 4,606 molecules and the remaining 979 molecules were unique to CV2-VWC1. Using only these 979 small molecules a mini library was and one LUDI search was performed on the entire BMP-2 binding region of CV2 using a 15 Å sphere of radius around the region. With the advanced LUDI parameters, we found 967 small molecules with Ludi3 binding scores greater than 300. We identified high scoring molecules that were distributed over the entire binding region and some of them are given in Table IV.4 and few docking poses are shown in Fig. 4.7.

3.4.5. Screening CV2-VWC1 Surface Using the ZINC and the CAP Databases and the Glide Docking

Single HTVS searches were performed for the selected residues of CV2-VWC1, BMP-2, BMPRIA, and BMPRII binding regions. In order to define the active site for receptor grid generations, we selected the centroid of the residues in the binding regions of CV2, BMP-2, BMPRIA, and BMPRII in the corresponding searches. Both the CAP and the ZINC databases were used separately with same parameters that were used

previously. For HTVS, we kept the default option of best-binding poses to be 10,000 and for SP and XP mode; we selected the best 1,000 high scoring molecules.

For CV2-VWC1 binding surface we performed Glide search in HTVS mode in the ZINC database and obtained a set of 10,000 high scoring molecules. Similarly, for the binding sites of BMP-2, BMPRIA, and BMPRII, we performed Glide-HTVS and obtained high scoring 10,000 molecules for each search. We sorted the 27,879 unique molecules among BMP-2, BMPRIA, and BMPRII and 1,927 common molecules among CV2-VWC1, BMP-2, BMPRIA, and BMPRII. These 1,927 small molecules were removed from the total unique CV2 blockers so that the small molecules that do not bind to BMP-2 and its receptors but only binds to the inhibitor, CV2 were used for further Glide searches. We found 8,072 exclusive CV2-VWC1 blockers from the ZINC database and with these molecules we performed Glide-SP and took the top scoring 1,000 molecules to use for Glide-XP search to get the final XP scoring.

We executed essentially the same procedure for the CAP database as described before. The CAP molecules found against CV2-VWC1, BMP-2, BMPRIA, and BMPRII were 10,000 for each protein. The unique molecules among BMP-2, BMPRIA, and BMPRII were 22,220 and the common molecules among CV2-VWC1, BMP-2, BMPRIA, and BMPRII were 5,851. After excluding these 5,851 molecules from unique CV2 blockers, we found 4,148 exclusive CV2 blockers from the CAP database and performed Glide-SP and XP mode docking and obtained XP scores for the top 1000 molecules.

3.4.6. Screening CV2-VWC1 Surface Using the ZINC and the CAP Databases and the GOLD Docking

The top scoring 1000 molecules were used each from the ZINC and the CAP databases from the above described Glide runs and performed GOLD docking. The PDB input files of single chains of the BMP-2, BMPRIA, BMPRII, and CV2-VWC1 were used here. We used the parameters that were used in BMP/noggin studies, and the GOLD searches were performed in background mode to obtain GOLD scores.

3.4.7. Consensus Scores for High Scoring Molecules from the Glide and the GOLD Docking on CV2-VWC1 Binding Surface

The consensus scores were calculated as CombScore from the Glide and the GOLD scores as described earlier. We then chose a set of molecules with high Glide, GOLD, and CombScore that were distributed over the hotspot binding region and some of these molecules are given in Table IV.5 and IV.6. In Fig. 4.8 and 4.9, we portrayed binding poses of some of these high scoring molecules. The compound VZGC001 with CombScore of 118, shown in Fig. 4.8, contains a hydrophilic region forming four hydrogen bonds and the hydrophobic region is pocketed into the hydrophobic grove in the receptor, CV2-VWC1, offering highest CombScore. Another molecule, VZGC003 with CombScore of 116, contains hydrophobic and hydrophilic aromatic groups, binds at almost similar location in a different orientation showing both the hydrophobic and hydrophilic regions pocketed into the respective complementary grooves with four hydrogen bonds. In Fig. 4.9, we have shown four other high scoring molecules from the CAP database in different locations on the binding hotspot region where they were tightly

docked into the grooves. In Fig. 4.10, we have shown the amino acids, T5, A7, I21, I27 and A36 whose mutations showed lower affinity in binding to BMP-2 [9] and in Fig. 4.6, 4.7, 4.8 and 4.9, we portrayed poses of a few of small molecules from the CAP and the ZINC database with high Ludi3 scores. Finally we have calculated the binding energy of all the selected molecules using MM-PBSA approach and the resulting values are given in Table IV.7.

3.5 Modeling and Design of Peptido-mimetic Compounds to Block Noggin/BMP Interaction to Promote Osteogenesis

Designing peptido-mimetic compounds is an alternative approach in developing potential drugs in blocking BMP/noggin interaction to promote osteogenesis instead of using small molecular databases. In our study, we identified several, di-, tri- and tetra-peptides derived from the key binding regions of BMP-2 that are in contact with noggin based on the analyses on the BMP-2/noggin complex structure. We modeled structures of these peptide sequences and performed some preliminary experiments using the Glide and the GOLD to dock these peptides onto respective binding sites on the noggin surface. We analyzed the scores from different docking procedures and compared the binding residues of noggin to BMP-2 protein. This study is simply the initiative for composing modified peptide mimetics to accomplish the proposed ultimate goal.

3.5.1 Identification and Selection of Di-, Tri- and Tetra-Peptides of BMP-2

Using the model complex of BMP-2/noggin, we calculated the PSA of each residue of noggin interacting with BMP-2 and *vice versa.* In Table I.2 we summarized the

residues of BMP-2 involved in binding with noggin residues. Depending on the BMP-2 binding epitope for the noggin, we identified five major tri- and tetra-peptides by analyzing the positions of the peptide residues in the model complex. Within these peptide sequences we identified subsequences relative to their positions with noggin. There were twenty one di- and tri-peptides originated from those five peptides. Altogether we used 26 peptides showed in the column 1 of Table V.1. The column 2 and column 3 of Table V.1 represent the binding residues of BMP-2 to noggin and *vice versa*.

3.5.2 Modeling and Making Peptide Small Molecular Database for Docking

While designing, we chose the BMP-2 peptides on the basis of their binding relationship with noggin residues. Only the regions of BMP-2 that were involved in binding to noggin were used to construct the peptides (Table V.1). We used Biopolymer module from Accelrys: Insight II and Build panel from Schrodinger's Maestro visualizer to create all these peptides from residue-by-residue. After that we optimized the geometry of these peptides in Maestro. Next, the Impact energy minimization was employed with default parameters to set up and perform energy minimization calculations on all these newly built peptides. Impact energy minimized structure files were used for each peptide to generate a peptide-mimetic database to dock onto noggin surface. This database was used for docking experiments to evaluate the ability of each peptide to dock at its corresponding binding location.

3.5.3 Docking the Peptide Molecules onto Their Respective Binding Sites

To predict and reproduce the correct binding modes and to examine the possible orientations of these peptides to the target protein, noggin, we exploited two computational docking programs that outperformed many other programs currently on the market [93, 94]. These are Schrodinger's Glide and CCDC's GOLD that are renowned for their aptitude to accurately mimic the experimental binding modes and excellence in scoring. These docking tools are very effective for computational simulation of the candidate ligand molecules binding to the receptor to predict the affinity and activity of binding molecules [93, 94].

We performed Glide in XP mode mostly using default parameters for accurate and advanced scoring of our peptides. First, we selected only the single chain of noggin protein from the complex PDB structure of 1m4u and then refined it in the Protein Preparation panel using 'Refinement only' on Maestro visualizer. Subsequently, we calculated receptor grids for noggin using Receptor Grid Generation panel. To define the active site we selected the centroid of the respective residues (Column 3 of Table V.1) where each peptide is expected to bind. Finally in the Ligand Docking panel we chose the XP mode to dock our peptides in a one-on-one fashion. We ensured that the 10 best-binding poses per peptide-ligand were selected to be written to the pose viewer file.

Next we executed the docking process using GOLD to authenticate the results derived from Glide docking. We started with loading the PDB input file of single chain of noggin on the GOLD Setup window and adding missing hydrogen atoms for protonation. After that we defined the binding site from a list of respective binding residues from noggin chain for each peptide ligand in separate text files. Then we selected the ligands

by specifying the peptides for corresponding searches and kept the default GA Runs as 10 per docking process. The maestro file format of each peptide ligand obtained from the energy minimization processes was converted to SD format to make it acceptable for GOLD runs. Later we chose the GoldScore as fitness function and performed GOLD by keeping all other parameters same as default. We accumulated the best scores from the summary of the fitness scores for all the docking attempts using our peptide database except for a few peptides. The text file comprising of the entire list of residues provided all solvent-accessible protein acceptor and donor atoms to the peptide based on their availabilities. We could not proceed with a few GOLD searches, such as for pep1.1, pep2.2, pep 3.2 and pep4.2 because the text file containing the binding residues did not have any required residues that can explicitly define the protein active site.

3.5.4 *Analysis of Docked Structures*

In this study, our aim was to design modified peptide mimetic compounds that potentially bind to noggin by mimicking the BMP-2 binding regions for noggin that eventually may prevent the binding of BMP-2/noggin leading to improved osteogenesis. Our present study is just the starting point of formulating and constructing modified peptide mimetics to fulfill such goal. Here we showed some preliminary docking experiments to gather some peptides with their scores to rationalize our plan. Table V.1 shows the Glide scores, (GlS in Column 4) and GOLD scores (GoS in Column 5) for the minimized peptides. There is still need to optimize these peptides and improve their scores from Glide and GOLD docking experiments. We compared Glide scores with the GOLD scores to check which peptides were the best in binding to the noggin residues.

This work helped us gain insights into modifying them to drug-like compounds by adding appropriate synthetically fusible side groups. The peptides with more negative Glide scores and more positive GOLD scores were indications of good results and form the basis for future modifications in their structure to better fit on the noggin surface. Figure 5.1 shows the binding poses of some of these peptides. Some of them contain hydrophilic regions forming up to five hydrogen bonds and their hydrophobic region is pocketed into the hydrophobic grove on noggin surface. To further test the binding mode of these peptides, we computed the PSA of each peptide residues in binding with noggin using the docking output files. In Table V.1, we summarized the residues of noggin that were in close contact with each peptide. This analysis is very critical in comparing the binding region of the noggin to each peptide with the original BMP-2/noggin complex.

4.0 DISCUSSION

The crystal structure of the BMP-7/noggin complex, the BMP-2/BMPRIA ectodomain complex and the extracellular domain of BMPRII monomer are known facilitating the analyses on BMP/noggin interaction. To review the activin/follistatin interaction, the crystal structure of the complex between activin ßA/follistatin and activin/ActRIIB extracellular domain are analyzed. The crystal structure of the BMP-2/CV2-VWC1 complex, the BMP-2/BMPRIA ectodomain complex and the extracellular domain of BMPRII monomer are known for BMP-2/CV2 study. From the analyses of structures of these complexes we identified the key amino acids present in the BMP-2 binding region of noggin, activin binding region of follistatin, and BMP-2 binding region of CV2. We employed various bioinformatics and *in silico* drug design methods to screen these regions to identify small molecular weight compounds from the CAP and the ZINC databases. We used the LUDI *de novo* design method for screening the potential compounds that block noggin, follistatin, and CV2 from interacting with their respective TGF-β ligands. We identified the high scoring potential binding ligands along with their theoretical binding scores. We verified the LUDI results with other computational techniques such as Glide and GOLD. Thus we selected a manageable number of these molecules for future experimental *in vitro* binding assays with the purified recombinant protein. If successful, these molecules will be taken for subsequent studies on BMP and BMP-like ligands-induced bone formation to test the BMP-potentiating effect by these compounds.

During the BMP/noggin study we found that BMP-2 binds an area of about 484 $Å^2$ on the noggin surface leading to its tight binding. A low molecular weight molecule most likely exhibit a high diffusion probability compared to a large BMP-2 molecule in binding noggin. We therefore feel that the advantage of designing tight binding small molecule will assist in competing with BMP-2 in binding to noggin competitively. From previous site directed mutagenetic studies it was observed that the mutants L46D, E48K, I218E of noggin fail to bind to BMP-2 [7]. This shows that any small changes on the surface of the noggin binding region renders the noggin surface lose binding to BMP-2 to bind, and this is what we are exactly trying to accomplish by targeting a tight binding small molecule onto the noggin, follistatin or on CV2-VWC1 before BMP-2 or any other TGF-β ligand recognizes the antagonist. We therefore identified a few high scoring potential binding molecules in the region surrounding these important antagonist binding surface residues that may block the binding to their respective ligands. In the case of noggin, Figs. 2.1, 2.2, 2.3 and 2.4 show binding poses of a few high scoring molecules and some of them block residues L46, E48, and I218.

From the activin/follistatin complex structure we estimated that activin binds to a total of 508 $Å^2$ area on the follistatin surface that provides compact and rigid binding conferring the binding of small molecules to follistatin very competitively when compared with the activin binding to follistatin. From previous site directed mutagenetic study we observed that the R192A mutant of follistatin fail to bind activin [51] proving this site to be a very good target site for small molecules. Taking this into account, our attempt was to target a tight binding small molecule onto this hotspot region before activin recognizes the follistatin. Therefore, we accumulated a few high scoring

molecules in the region surrounding this important residue with a desire to block this residue from binding with activin. The poses of some of these molecules are shown in Fig. 3.6, 3.7, 3.8 and 3.9.

While monitoring the BMP-2/CV2 interaction, we observed that BMP-2 binds to an area of about 271 $Å^2$ on the CV2-VWC1 surface leading to a constructive binding. A snug fitting small molecule is more likely to compete with BMP in binding to VWC1 because of its tendency of high diffusion during VWC1 binding. Previous truncation and mutagenetic studies endorsed that the T3P, T5P, 121A, I21R, A36R and A7 truncation mutants of CV2 exhibit significant lower affinity to bind to BMP-2 [9]. This explains that any alteration on the surface of the CV2-VWC1 binding region regarding these residues results in loss of BMP-binding to VWC1 surface. This scenario fulfills our goal since we would like the small molecules to bind onto the VWC1 before BMP-2 interacts with the VWC1 domain of CV2. As a result, we have gathered a number of high ranking molecules in the region neighboring these key residues looking forward to preventing BMP-2 binding to these residues and Figs. 4.6, 4.7, 4.8 and 4.9 display binding poses of some of these high scoring molecules surrounding these residues.

In our BMP/noggin study, we proposed another strategy of using peptides built from the key binding residues of BMP-2 sequence that are in contact with noggin based on the BMP-2/noggin complex structure and then modeled several, di-, tri- and tetra-peptide combination structures. Then with the help of two molecular docking techniques, we docked these peptides onto respective binding sites on noggin and analyzed the scores from two different techniques. Fig 5.1 portrays some of these high scoring peptides from

our constructed peptide database. This is an alternative strategy in blocking BMP/noggin binding to promote osteogenesis instead of using the commercially available small molecular databases.

Each of the computational technique operates by assigning varying approaches and scoring schemes that the developers followed. Several previous method evaluation studies have shown that different methods have succeeded for different combination of receptor ligand targets [94, 95]. Priorities for particular approach can limit the expectations in different techniques. A very good criteria specific for one technique could be a constraint for the others. Different computational studies are therefore desirable to acquire comparatively matching outcome. We therefore performed the docking using the three major successful methodologies, GOLD, Glide, and LUDI to authenticate the results obtained and to select a manageable number of high scoring compounds for future experimental verification. Thus, this study provides initial attempts towards finding the new low molecular weight and drug-like compounds to augment the effects of BMP and BMP-like compounds in bone induction and to consequently reduce their required dosage and cost, although more definitive biological studies will be required to validate the strategies employed in the current study.

5.0 FUTURE DIRECTIONS

The greatest advantage of computational screening of small molecules in drug-designing is to compute reliable scores with defined protein-ligand binding affinity calculations. This exercise results in screening the millions of available molecules to a few tens or hundreds which can be easily tested experimentally to determine a lead molecule. However, a lack of universal and precise scoring method to determine the actual binding affinity is still a far cry. Here we present our computational analysis and different molecular docking studies to nominate for prospective molecules that bind the antagonists such as noggin, follistatin, and CV2 from the available drug-like small molecular databases and from our own database of peptides. In future our upcoming projects will be to test their binding affinity and biological efficacy experimentally for all the scoring molecules using biochemical and cell-based assays which we already undertook with a joint collaboration with organic chemists, biochemists, and orthopedic surgeons.

6.0 TABLES

Table I.1. Amino acids of Noggin, BMPRIA, BMPRII and BMP-2 involved in binding**.** Surface accessible contact area in Å (sum) and percentage of total residue area (% SA) involved in binding is given. Only the residues that show a loss of greater than 10% SA upon contact are given.

Noggin (1m4u)				BMPRIA (2h62)				BMP-2 With					
										Receptors		Noggin	
Pos	AA	Sum	%SA	Pos	AA	Sum	%SA	Pos	AA	Sum	%SA	Sum	%SA
27	MET	7.0	11.3	42	GLY	10.1	42.5	10	LEU	29.2	51.7	-	-
30	TYR	20.5	33.0	43	HIS	22.1	40.2	25	ASP	9.0	22.8	-	-
31	LEU	11.2	19.7	45	PRO	13.1	29.4	26	VAL	7.3	15.4	-	-
32	HIS	30.1	54.7	46	ASP	4.0	10.1	27	GLY	4.7	19.5	-	-
33	ILE	20.3	36.6	67	ASP	11.8	30.0	28	TRP	14.8	19.6	9.3	12.4
34	ARG	37.9	52.2	72	THR	4.6	11.0	29	ASN	-	-	9.1	22.3
35	PRO	39.6	88.7	77	CYS	9.7	23.1	30	ASP	4.5	11.5	4.1	10.6
36	ALA	10.0	30.2	79	LYS	25.2	41.0	31	TRP	19.5	25.9	20.1	26.7
37	PRO	24.1	54.2	81	GLU	7.9	16.2	33	VAL	14.8	31.0	15.2	31.8
38	SER	11.6	34.2	82	GLY	10.9	46.0	34	ALA	12.2	36.7	12.1	36.3
39	ASP	8.6	21.7	84	ASP	20.6	52.4	35	PRO	8.8	19.6	9.6	21.4
40	ASN	16.1	39.3	85	PHE	51.3	84.4	36	PRO	-	-	15.1	33.8
43	LEU	23.5	41.6	86	GLN	28.6	55.4	39	HIS	18.5	33.7	-	-
46	LEU	32.9	58.1	88	LYS	27.4	44.6	49	PHE	29.7	48.9	10.5	17.3
47	ILE	8.9	16.1	89	ASP	9.8	24.7	50	PRO	30.7	69.0	17.1	38.3
49	HIS	7.1	12.9	90	SER	17.6	52.1	51	LEU	6.2	11.1	7.1	12.6
199	HIS	11.2	20.4	91	PRO	5.1	11.3	52	ALA	11.5	34.5	9.2	27.7
204	ARG	11.2	15.4	92	LYS	26.1	42.6	53	ASP	36.0	91.4	21.6	54.7
206	ARG	27.8	38.4	93	ALA	3.6	10.9	54	HIS	27.4	49.8	-	-
208	GLN	7.3	14.2	94	GLN	23.6	45.8	57	SER	8.2	24.3	-	-
209	ARG	18.5	25.5	97	ARG	15.8	21.8	59	ASN	7.2	17.6	9.3	22.6
210	ARG	31.2	43.0	**BMPRII (2hlr)**				62	ILE	20.15	36.3	21.94	20.2
218	ILE	14.1	25.4	Pos	AA	Sum	%SA	66	LEU	13.5	23.9	-	-
219	PRO	13.9	31.2	40	TYR	23.3	37.5	69	SER	18.1	53.6	-	-
220	ILE	12.6	22.7	67	TYR	6.5	10.4	70	VAL	19.3	40.4	-	-
221	GLN	22.0	42.7	69	LEU	11.7	20.7	72	SER	6.2	18.4	-	-
223	PRO	5.3	12.0	81	LYS	15.2	24.8	76	LYS	12.3	19.9	-	-
				84	CYS	8.8	21.0	85	SER	7.3	21.6	-	-
				85	TRP	20.8	27.7	86	ALA	7.8	23.6	-	-
				86	SER	15.3	45.4	87	ILE	9.7	17.5	-	-
				87	HIS	18.9	34.3	88	SER	15.8	46.8	14.8	43.8
				88	ILE	19.1	34.5	90	LEU	17.2	30.5	19.2	33.9
				89	GLY	12.2	51.4	92	LEU	-	-	9.9	17.5
				90	ASP	15.5	39.5	96	GLU	9.1	18.8	-	-
				93	GLU	7.9	16.3	97	LYS	13.8	22.5	17.9	29.1
				113	TYR	16.4	26.3	98	VAL	21.5	45.1	29.7	62.2
				115	PHE	13.0	21.4	99	VAL	-	-	10.4	21.8
								100	LEU	23.1	37.6	36.2	64.0
								101	LYS	11.5	18.7	23.8	38.7
								102	ASN	6.3	16.0	20.2	49.5
								103	TYR	10.1	16.2	7.6	12.3
								104	GLN	-	-	11.0	21.2
								107	VAL	4.8	10.1	-	-

Table I.2. Amino acids of BMP-2 involved in binding to receptors BMPR-IA, BMPR-II, and noggin are identified using the modeled protein complexes of BMP-2/BMPR-IA/BMPR-II and BMP-2/noggin.

BMP-2		BMPRIA	BMPRII	Noggin
10	LEU	42-43,67,72		
25	ASP	90-92		
26	VAL	85,88-93		34
27	GLY	85,88-92		209-210
28	TRP	85,88-91		34,208-210
29	ASN			206,208-210
30	ASP	88		208-210
31	TRP	84-85,88		34,208-209
33	VAL		69,113,115	46,49,204,206,208-210
34	ALA		67,69,85,93,115	46-47,49,204,206,210
35	PRO		85,87,93,115	46-47,49,210
36	PRO		93	46-47,49,206,209-210
39	HIS		87-90,93	47
49	PHE	43,86,89,97		31-33
50	PRO	43,77,82,85-86,89		27,30-34
51	LEU	43,77,82,85-86		30-34
52	ALA	42-43,45,77,79,82,86		27,30-33
53	ASP	43,45,77,79,81-82		27,30-32
54	HIS	43,45-46,77,79		27,30
57	SER	79,81-82		32,35-36
59	ASN	79,81-82,84-85		32-37
62	ILE	82,84-86,89		31-36
66	LEU	85-86,89-90,92-94		32-34
69	SER	89-94		
70	VAL	89-94		
72	SER	93-94		
85	SER		86-89	
86	ALA		85-89	37-38,40,43
87	ILE		85-90	37-38,40,43,46-47
88	SER		67,84-88	37-40,43,46-47
90	LEU		67,69,81,85,115	37,46-47,204,206,220-221
92	LEU			206,218-221
96	GLU		40	206,208,218-220
97	LYS		40,81	204,206,218-221
98	VAL		40,69,81	204,206,218-221
99	VAL		40,81	37,199,204,219-221,223
100	LEU		40,81,84-85	37,39,43,46,199,220-221,223
101	LYS	81	84-86	36-40,43,46,199,221,223
102	ASN	81	84-87	36-40,43,46,223
103	TYR	81-82	86	35-40,43
104	GLN	81-82	86	35-40
107	VAL		87-90	

Table I.3. The LUDI virtual screening advanced input parameters used for docking runs.

	Input Parameter Name	Value Used
Fragment Filters	Maximum Atoms in Fragment	300
Fitting	Maximum Fit Attempts	9999
	Maximum RMSD	0.4
	Bond Rotation	One at a time
Hit Criteria	Minimum Surface	10
	Minimum Score	300

Table II.1. Some selected top ranking representative noggin ligands in the hotspot regions from the CAP database using the LUDI docking runs. Absolute energy (ABS), percentage of fragment surface contact area with receptor (PCA), score from hydrogen bond term (HBS), lipophilic term (LPS), non hydrogen bonds (NHB), RMSD of compound fit and the total Ludi3 Score (L3S) are given. The percent solvent accessible contact area covered by each molecule on binding to noggin residues is also given in parenthesis.

Molecules	ABS	PCA	HBS	LPS	NHB	RMSD	L3S	Noggin AAs in contact with the CAP small molecule (in % area)
VCLN001	51.2	19	145	56	3	0.43	558	P37(22.2), A36(17.4), P35(1.2), R34(27.3)
VCLN002	95.9	29	47	95	1	0.32	551	P223(4.6), Q221(10), H199(30.7), D39(40.4), P37(20.6)
VCLN003	52.8	16	55	33	1	0.40	503	P37(2.7), A36(45.4), P35(10.5), H32(34.9), M27(15.1)
VCLN004	50.9	46	96	103	2	0.41	479	P223(6.7), H199(29), D39(45.4), S38(1.7)
VCLN005	77.5	17	0	35	0	0.34	462	Q221(16.1), H199(18.1), D39(32.4), S38(3.1), P37(10.7)
VCLN006	31.6	36	55	101	1	0.37	461	P223(3.2), Q221(21.2), H199(25.6), D39(41.7)
VCLN007	95.2	34	100	89	2	0.40	459	P223(1.3), Q221(16.2), H199(26.8), D39(31.1),
VCLN008	90.9	30	97	106	3	0.42	457	P223(17.9), Q221(24.9), I220(9.5), H199(18.5), L43(9.4), D39(27.2), S38(7)
VCLN009	31.6	37	92	81	3	0.59	443	P37(26.6), A36(29.6), P35(2.8), R34(10.7), H32(9.8)
VCLN010	67.6	20	153	29	4	0.54	443	A36(13.4), H32(20.3), M27(31.3)
VCLN011	26.3	37	101	89	2	0.35	440	P223(4.2), Q221(4.9), H199(24.5), D39(39)
VCLN012	43.5	32	91	41	2	0.44	437	P223(0.7), Q221(4.1), H199(21.9), D39(33.6)
VCLN013	56.8	52	54	158	1	0.37	427	P223(13.8), H199(30.2), L43(6.8), D39(43.5), S38(6.9)
VCLN014	158.0	30	64	68	2	0.42	402	P37(30.8), A36(29.1), R34(9.5), H32(7.9), L31(1.8)
VCLN015	145.7	30	166	74	4	0.24	400	P223(13.6), Q221(11.8), I220(7.2), H199(4.4), L46(13.6), L43(17.8), D39(0.5), S38(2.8), P37(1)
VCLN016	32.2	50	73	168	2	0.35	550	P223(6.6), Q221(13.7), I220(21.1), R204(5.6), L46(17.4), L43(8.7)
VCLN017	50.4	41	55	133	1	0.33	541	R210(1.5), Q208(4.5), R206(12.7), R204(9.1)
VCLN018	183.2	10	42	22	1	0.35	522	R210(9), E48(3.6), I47(18.6), L46(9.5)
VCLN019	24.0	46	101	124	3	0.43	515	R206(3.6), R204(13.5), E48(5.3), I47(4.4), L46(4.7)
VCLN020	39.1	47	42	155	2	0.35	506	R206(1.9), R204(6.5), E48(4.4), I47(12), L46(0.5)
VCLN021	53.1	30	110	66	2	0.35	485	R210(13.5), Q208(3.5), R204(6), E48(4.8), I47(2)
VCLN022	44.1	52	86	120	2	0.32	476	R204(4.6), E48(5.2), I47(16.7), L46(1.1)
VCLN023	22.7	38	154	120	3	0.42	473	R210(1.2), R206(4.1), R204(10.1), E48(5), I47(1.1)
VCLN024	70.9	46	78	158	2	0.30	451	R206(6.7), R204(13.1), E48(5.3), I47(1.8), L46(5.7)
VCLN025	21.0	42	62	103	2	0.37	447	R206(3.7), R204(13.5), E48(5.3), I47(4.6), L46(4.8)
VCLN026	177.3	33	55	116	1	0.37	441	R210(6.8), Q208(1.9), R204(6.3), E48(5.1), I47(3)
VCLN027	47.2	34	110	106	2	0.4	431	R206(4.7), R204(11.4), E48(4.5), I47(1.2)
VCLN028	29.9	48	130	141	4	0.41	431	R206(4.7), R204(8.5), E48(4), I47(2)
VCLN029	16.0	30	48	74	1	0.35	430	Q208(0.5), R206(12.7), R204(11.7), E48(2.3)
VCLN030	92.0	31	104	54	4	0.59	424	R210(6.2), Q208(10.1), R206(9.7), R204(7.7)
VCLN031	31.8	41	84	128	2	0.42	424	R210(1), Q208(2), R206(5), R204(10.7), E48(3.2)

Table II.2. Some selected top ranked representative noggin ligands in the hotspot regions from the ZINC database using the LUDI docking runs. Absolute energy (ABS), percentage of fragment surface contact area with receptor (PCA), score from hydrogen bond term (HBS), lipophilic term (LPS), non hydrogen bonds (NB), RMSD of the compound fit and total Ludi3 Score (L3S) are given. The percentage solvent accessible contact area covered by each molecule on binding to the noggin residues is also given in parenthesis.

Molecules	ABS	PCA	HBS	LPS	NB	RMSD	L3S	Noggin AAs in contact with small molecules (in % area)
VZLN001	48.6	44	48	205	1	0.34	562	L43(15.9), L46(24.6), H199(2.6), R204(2.9), I220(11.4), Q221(11.2), P223(10.4)
VZLN002	45.1	38	76	187	2	0.43	556	D39(4.5), L43(5.3), L46(14.3), H199(14.2), R204(3.3), I220(19.4), Q221(27.5), P223(8.2)
VZLN003	116.2	50	59	172	2	0.49	556	L43(8.6), L46(17), R204(3.1), I220(17.5), Q221(15.3), P223(6.3)
VZLN004	103.8	36	100	153	2	0.37	523	A36(0.1), P37(16.4), S38(15.7), L43(24.4), L46(21.4), I220(5.1), Q221(2.4), P223(4.7)
VZLN005	64.5	42	236	24	6	0.36	514	I47(2), E48(2.7), H49(14.1), R204(8.7), Q208(9.8), R210(4.2)
VZLN006	122.9	13	55	24	1	0.24	510	P37(8.3), D39(40.8), H199(27.3), Q221(17.8)
VZLN007	33.5	42	146	147	3	0.34	508	P37(1.8), S38(10.3), D39(13.6), N40(0.2), L43(19.2), L46(11.7), H199(8.6),I220(5.5), Q221(7.7), P223(18)
VZLN008	97.8	22	47	47	1	0.32	503	P37(16.8), D39(0.3), D39(40.2), H199(30.4), Q221(8.6)
VZLN009	37.2	49	110	66	2	0.14	501	M27(3.8), H32(15.9), R34(12), P35(3), A36(40.3), P37(14.4)
VZLN010	33.7	34	99	180	2	0.37	501	P37(8.3), S38(0.1), D39(33.4), L43(8.9), L46(9.3), H199(20.3), I220(13.4), Q221(21), P223(8.9)
VZLN011	117.2	37	110	172	2	0.24	497	P37(21.6), S38(17.5), D39(20.8), L43(24.8), L46(13.6), H199(16.3), I220(3.3), Q221(10.6), P223(14.8)
VZLN012	28.4	40	0	187	0	0.38	496	L43(18.5), L46(17.7), I220(15.3), Q221(6), P223(13.3)
VZLN013	56.5	38	110	60	2	0.13	495	M27(3.8), H32(15.9), R34(13.6), P35(3.1), A36(40.1), P37(18.9)
VZLN014	91.2	40	87	170	3	0.44	495	P37(8.3), S38(9.7), D39(21.7), L43(10.9), L46(17.5), H199(18.6), I220(7.9), Q221(16), P223(15.3)
VZLN015	59.0	75	103	10	2	0.39	493	I47(8.3), E48(4.4), H49(25.1), R204(3.6)
VZLN016	85.5	47	0	151	0	0.43	492	P37(18.8), S38(9.5), D39(35.5), L43(5), H199(26.1), P223(9.9)
VZLN017	45.7	44	49	170	1	0.44	489	L43(11.6), L46(27.3), R204(7.7), I218(2.5), I220(14.6)
VZLN018	49.3	43	93	70	2	0.32	488	M27(1.2), H32(8.9), R34(1.8), P35(2.5), A36(28), P37(16.8)
VZLN019	37.3	33	101	180	2	0.26	487	P37(17), S38(9), D39(13.5), L43(13.6), L46(13.5), H199(10.8), I220(7.8), Q221(19.6), P223(16.2)
VZLN020	26.8	35	110	164	2	0.1	480	P37(7.5), S38(0.8), D39(23.9), L43(11.5), L46(12.6), H199(19.6), I220(9.2), Q221(21.1), P223(13.1)
VZLN021	61.2	35	45	64	1	0.41	479	L46(2.2), E48(5), H49(3.9), R204(13.3), Q208(2.4)
VZLN022	101.6	41	95	124	2	0.45	473	L43(12.4), L46(24.6), R204(2.4), I220(15), P223(3.6)
VZLN023	152.1	32	110	37	2	0.3	472	I47(3.8), E48(5.1), H49(12.9), R204(10.9)
VZLN024	80.0	33	89	183	2	0.44	471	L43(1.9), L46(12.5), R204(6), I218(6.9), P219(0.7), I220(24.3), Q221(24.5), P223(4.8)
VZLN025	59.9	35	101	170	2	0.37	470	L43(7.7), L46(26.6), R204(7), I218(6.8), P219(0.7), I220(22.4), Q221(13.7), P223(6.1)
VZLN026	82.0	46	19	180	1	0.46	469	L43(9.4), L46(16.8), I220(17.9), Q221(15.8), P223(6.9)
VZLN027	87.5	29	108	162	2	0.29	469	L43(9.4), L46(20.6), H199(8.8), I220(13.1), Q221(21.8), P223(6)
VZLN028	21.9	44	83	114	2	0.35	467	L43(14.9), L46(20.2), R204(0.5)I220(9), Q221(3.1), P223(4.9)
VZLN029	41.0	43	0	228	0	0.34	466	S38(6.3), D39(20.2), L43(14), L46(10), H199(13.5), I220(8.7), Q221(13), P223(18.7)
VZLN030	27.0	44	52	143	2	0.44	465	L43(13), L46(28.1), R204(1.1), I220(10.8), Q221(2.3), P223(2.8)

Table II.3. Some selected top scoring representative noggin ligands from the CAP database using the Glide-XP and the GOLD docking runs. Also shown, are some selected top ranking molecules using combined, Glide-XP (GlS) and GOLD (GoS) and CombScore (CoS) scores in the hotspot region of noggin. The solvent accessible contact area covered by each small molecule on binding to noggin residues is also given in parenthesis. The scores considered in selecting the respective representative molecules from the group are shown in **bold** letters.

Molecules	GlS	GoS	CoS	Noggin AAs in contact with small molecules (in % area)
VCGN001	-9.1	54.9	**122.6**	P223(4.5), Q221(10.6), I220(22.7), P219(2.5), I218(11.2), R206(2.2), R204(6.1), L46(21.7), L43(12.9)
VCGN002	-9.0	55.1	**121.9**	P223(4.8), Q221(10.6), I220(22.7), P219(3.9), I218(11.1), R206(6.4), R204(6.5), L46(19.1), L43(11.2)
VCGN003	-8.3	56.2	**117.7**	P223(6.9), Q221(20.7), I220(20.6), P219(3), R204(2.5), L46(15.8), L43(11.8)
VCGN004	-9.1	49.2	**116.9**	P223(5.1), Q221(6), I220(22.7), I218(7.9), R204(7.6), L46(19), L43(10.9)
VCGN005	-9.2	47.5	**115.8**	P223(7.6), Q221(15.3), I220(19.6), R204(4), L46(24.5), L43(17.8)
VCGN006	-8.1	55.2	**115.4**	P223(10), Q221(26.3), I220(22.9), P219(5.3), I218(6.5), L46(14.7), L43(13.5)
VCGN007	-9.5	44.7	**115.0**	P223(17.6), Q221(9.7), I220(9.7), H199(8), L46(16.2), L43(24.5), D39(14.7), S38(13.2), P37(6.3)
VCGN008	-8.1	53.6	**113.7**	P223(4.9), Q221(11), I220(22.7), I218(13.3), R206(5.5), R204(5.9), L46(19.1), L43(11.6)
VCGN009	-10.3	36.6	**113.0**	P223(13.1), Q221(6.3), I220(9.2), L46(20), L43(26), S38(22.7), P37(14.5)
VCGN010	-8.6	48.8	**112.9**	P223(4.9), I220(19.4), I218(6.9), R206(5.1), R204(10.7), L46(20), L43(10.9)
VCGN011	-4.5	**62.3**	95.9	P223(17.6), Q221(18), I220(17.6), R204(4.6), H199(16.4), L46(16.5), L43(23.1), D39(23.9), S38(13.4), P37(14.7)
VCGN012	-5.6	**61.3**	102.8	Q221(6.7), I220(18.5), I218(3.6), R206(17.2), R204(16.9), L46(14.6)
VCGN013	-6.0	**61.1**	106.0	Q221(6.8), I220(21.5), I218(3.8), R206(7.8), R204(19.9), E48(1.2), L46(14.1)
VCGN014	-4.1	**59.2**	89.3	P223(6.1), Q221(8.2), I220(20.7), I218(8.7), R206(17.5), R204(15.5), L46(11), L43(4.6)
VCGN015	-6.9	**58.8**	109.9	P223(6.4), Q221(7.9), I220(15.8), I218(4.5), R206(11.1), R204(16.4), L46(22.2), L43(9.6)
VCGN016	-5.9	**58.8**	102.4	P223(5.7), Q221(7.5), I220(18.1), R206(9.9), R204(19.8), L46(18.8), L43(5.7)
VCGN017	-5.2	**58.6**	96.9	P223(4.6), Q221(5), I220(17.4), R206(9.9), R204(19.9), L46(32.7), L43(14.8)
VCGN018	-5.6	**58.0**	99.5	Q221(8.6), I220(20.2), I218(6), R206(11), R204(17), L46(12.9)
VCGN019	-4.8	**57.9**	93.8	Q221(4.3), I220(23.1), P219(4.4), I218(12.7), R204(7.1), L46(25.5), L43(15.5)
VCGN020	-5.7	**57.8**	99.8	P223(4.2), Q221(3.1), I220(11), R206(10.6), R204(17.8), L46(26.2), L43(4.3)
VCGN021	**-9.9**	32.4	105.6	P223(5.1), Q221(10.1), I220(21.9), I218(9), R204(8.5), L46(20), L43(10.8)
VCGN022	**-9.4**	35.5	105.1	P223(11.2), Q221(9.9), I220(5.7), H199(3.3), L46(21.6), L43(25.1), P37(5.8)
VCGN023	**-9.3**	42.6	111.5	P223(5.5), Q221(10.8), I220(22), I218(5.1), R204(10.6), L46(25.6), L43(11.7)
VCGN024	**-8.9**	40.4	106.2	P223(5.1), Q221(9.5), I220(20.6), R204(13.2), L46(23.9), L43(11.6)
VCGN025	**-8.8**	43.3	108.6	R210(12.1), Q208(5.1), R206(2.3), R204(7.2), E48(4.3)
VCGN026	**-8.8**	37.0	102.3	P223(9.4), Q221(21.4), I220(20.9), P219(1.8), L46(17.8), L43(13.8)
VCGN027	**-8.8**	40.3	105.5	P223(14.6), Q221(19.5), I220(16), H199(12.9), L46(16.8), L43(14.9), D39(12.6)
VCGN028	**-8.7**	37.9	102.6	P223(9.9), Q221(17.6), I220(18.5), H199(5.6), L46(18.6), L43(14.9)
VCGN029	**-8.7**	48.3	112.8	P223(11.2), Q221(21.5), I220(16.8), H199(6.4), L46(23.1), L43(15.8)
VCGN030	**-8.7**	45.4	110.0	P223(8.4), Q221(8.9), I220(10.8), R204(10.2), L46(34), L43(16.6)

Table II.4. Some selected top scoring representative noggin ligands from the ZINC database using the Glide-XP and the GOLD docking runs. Also shown, are some selected and top ranked molecules using combined Glide-XP (GlS), GOLD (GoS) and CombScore (CoS) scores in the hotspot region of noggin. The solvent accessible contact area covered by each small molecule on binding to noggin residues is also given in the parenthesis. The scores considered in selecting the representative molecules from the group are shown in **bold** letters.

Molecules	GlS	GoS	GoS	Noggin AAs in contact with small molecules (in % area)
VZGN001	-10.0	60.0	**139.5**	P223(9.5), Q221(6.6), I220(16.1), R204(3.5), L46(15.6), L43(25.4)
VZGN002	-10.2	50.6	**132.2**	P223(14.9), Q221(22.9), I220(18.2), P219(5.5), H199(7.3), L46(14.4), L43(17.7), D39(4.4)
VZGN003	-8.6	60.1	**128.7**	P223(5.7), Q221(9.3), I220(19), R206(9.5), R204(18), L46(19.5), L43(10.8)
VZGN004	-9.0	54.5	**126.1**	P37(16.5), A36(35.4), R34(23.1), I33(9.2), H32(10.2), L31(7.7), M27(5.7)
VZGN005	-9.0	52.9	**124.8**	P37(10.3), A36(36.1), R34(24.8), I33(9.9), H32(10.6), L31(6.7), M27(5.6)
VZGN006	-8.0	59.9	**123.7**	P223(5.5), Q221(22), I220(23.5), R204(5.9), L46(17.8), L43(10.2)
VZGN007	-8.7	53.8	**123.6**	Q221(12), I220(22.7), I218(8.7), R204(5.7), L46(17.7), L43(10.5)
VZGN008	-8.2	57.4	**122.8**	P223(4.9), I220(15.6), R206(11.4), R204(18.4), L46(25.8), L43(12.6)
VZGN009	-8.6	52.1	**120.9**	P223(8.1), Q221(6.7), I220(6), I47(11), L46(25), L43(27.9)
VZGN010	-8.0	56.8	**120.8**	Q221(1.8), I220(4.1), I47(11.1), L46(21.1), L43(25.1)
VZGN011	**-9.7**	41.8	119.1	P223(5.4), Q221(9.3), I220(19.8), I218(5.7), R204(10.4), L46(20), L43(10.4)
VZGN012	**-9.4**	45.2	120.4	P223(14), Q221(21.3), I220(14.9), H199(10.5), L46(16.6), L43(15.8), S38(5.8)
VZGN013	**-9.4**	41.7	116.6	Q221(12), I220(19.4), R206(4), R204(9.1), L46(20.3), L43(11.3)
VZGN014	**-9.1**	40.7	113.5	P223(16.6), Q221(9.7), I220(3.6), H199(12.3), L46(14.5), L43(24.7), D39(18.9), S38(15.3), P37(5.2)
VZGN015	**-9.1**	43.0	115.6	P37(6.2), A36(30.8), R34(22.7), I33(14.2), H32(10.3), L31(12.8)
VZGN016	**-9.1**	37.6	110.0	P223(9.8), Q221(16.9), I220(17.9), R204(8.8), L46(25.4), L43(16.5)
VZGN017	**-9.0**	33.0	104.9	P223(12.9), Q221(7.5), I220(18.6), R204(6.4), H199(1), L46(21.5), L43(19.8), S38(5.7), P37(4.7)
VZGN018	**-8.9**	46.2	117.5	Q221(9.1), I220(21.4), I218(9.3), R204(8.1), L46(19.7), L43(11.1)
VZGN019	**-8.9**	34.0	104.7	P223(5.6), Q221(5.1), I220(13.2), R204(10.6), L46(39.8), L43(14.5)
VZGN020	**9.0**	47.4	119.2	P223(5.5), Q221(9.6), I220(22.1), I218(8.2), R204(7), L46(17.8), L43(10.3)
VZGN021	-6.2	**63.5**	112.7	P37(36.2), A36(41.6), P35(10.7), R34(18.1), H32(14), M27(5.5)
VZGN022	-3.2	**63.1**	88.9	I220(9.4), I218(3.9), R206(17.9), R204(19.6), E48(4), L46(6.5)
VZGN023	-6.0	**62.7**	110.7	P223(5.8), Q221(5.6), I220(19.9), R206(10.7), R204(13.2), L46(18), L43(11)
VZGN024	-0.4	**61.8**	65.1	P223(9.8), Q221(10.9), I220(16.3), R204(19.9), L46(33.2), L43(12.3)
VZGN025	-4.2	**61.8**	95.2	Q221(5), I220(16.2), R206(8.7), R204(19.9), L46(31.3), L43(10)
VZGN026	-4.9	**61.6**	100.6	P223(5.6), Q221(7.7), I220(20.2), R206(10.6), R204(18.7), L46(17.1)
VZGN027	-5.7	**60.8**	106.6	I220(14.7), R206(9), R204(18), L46(29.4), L43(10.7)
VZGN028	-5.7	**59.9**	105.2	P223(5.9), I220(12.4), R206(10.8), R204(15), L46(25.5), L43(11.9)
VZGN029	-6.7	**59.9**	113.2	P37(10.7), A36(32.8), R34(18.6), I33(12.7), H32(10.1), L31(11.2)
VZGN030	-3.3	**59.2**	85.6	P223(9.7), Q221(15.6), I220(23.4), I218(6.2), R204(11.5), L46(22), L43(13.1)

Table II.5. Binding energy values calculated using Molecular Mechanics Poisson-Boltzmann Surface Area (MM-PBSA) method for the top scoring representative noggin ligands from the CAP and the ZINC databases given in Table II.3 and II.4.

CAP Molecules in Table II.3	**MM-PBSA Binding Energy**	**ZINC Molecules in Table II.4**	**MM-PBSA Binding Energy**
VCGN001	-31.05	VZGN001	-7.73
VCGN002	-19.62	VZGN002	-34.75
VCGN003	-29.31	VZGN003	-13.54
VCGN004	-16.93	VZGN004	-21.52
VCGN005	-24.77	VZGN005	-30.70
VCGN006	-27.19	VZGN006	-27.55
VCGN007	-18.12	VZGN007	-28.28
VCGN008	-22.60	VZGN008	-20.32
VCGN009	-21.14	VZGN009	-20.59
VCGN010	-17.60	VZGN010	-13.55
VCGN011	-29.97	VZGN011	-22.04
VCGN012	-7.52	VZGN012	-16.72
VCGN013	2.07	VZGN013	-24.22
VCGN014	-11.43	VZGN014	-11.65
VCGN015	-20.96	VZGN015	-8.26
VCGN016	-6.44	VZGN016	-25.05
VCGN017	-7.57	VZGN017	-34.48
VCGN018	-12.01	VZGN018	-15.51
VCGN019	-20.79	VZGN019	-13.50
VCGN020	-5.09	VZGN020	-22.02
VCGN021	-24.60	VZGN021	-22.13
VCGN022	-28.74	VZGN022	-9.22
VCGN023	-29.98	VZGN023	-25.51
VCGN024	-32.48	VZGN024	-17.69
VCGN025	-4.49	VZGN025	9.78
VCGN026	-30.79	VZGN026	-9.40
VCGN027	-31.87	VZGN027	-21.33
VCGN028	-23.60	VZGN028	-23.29
VCGN029	-29.83	VZGN029	-13.04
VCGN030	-21.29	VZGN030	-26.72

Table III.1. Amino acids of follistatin, ActRIIB and activin interacting region**.** Surface accessible contact area in Å (Sum) and percentage of total residue contact area (% SA) involved in binding are given.

Follistatin (2b0u)				**ActRIIB (1s4y)**				**Activin with**					
Pos	**AA**	**Sum**	**%SA**	**Pos**	**AA**	**Sum**	**%SA**	**Pos**	**AA**	**ActRIIB (1s4y)**		**Follistatin (2b0u)**	
										Sum	**%SA**	**Sum**	**%SA**
5	LEU	10.2	17.9	13	ASN	10.8	10.8	1	GLY	-	-	13.0	54.9
6	ARG	14.2	19.6	16	LEU	7.3	7.3	2	LEU	-	-	18.6	32.8
14	GLN	6.6	12.7	17	GLU	7.7	7.7	3	GLU	-	-	5.6	11.4
15	VAL	23.4	49.0	30	GLU	17.2	17.2	5	ASP	-	-	5.6	14.3
16	LEU	20.9	37.0	33	LYS	17.1	17.1	8	VAL	-	-	8.5	17.7
17	TYR	23.5	37.7	38	TYR	6.9	6.9	10	ILE	-	-	12.7	22.9
31	ARG	14.8	20.4	42	ARG	9.8	9.8	25	TRP	-	-	8.1	10.8
44	ASN	14.1	34.4	51	VAL	9.0	9.0	27	ASP	-	-	9.3	23.5
46	LEU	5.9	10.5	52	LYS	18.5	18.5	28	TRP	-	-	26.9	35.7
47	PHE	35.7	58.7	55	CYS	4.9	4.9	30	ILE	13.3	24.0	13.2	23.9
48	LYS	11.6	18.9	56	TRP	27.6	27.6	31	ALA	12.4	37.3	9.2	27.5
50	MET	14.3	23.2	57	LEU	36.5	36.5	32	PRO	11.3	25.4	10.1	22.7
51	ILE	33.2	59.7	58	ASP	11.2	11.2	33	SER	5.6	16.4	-	-
52	PHE	34.8	57.3	59	ASP	5.2	5.2	36	HIS	9.3	16.9	14.8	27.0
105	LEU	15.6	27.6	60	PHE	17.6	17.6	47	HIS	-	-	17.5	31.7
106	ASP	4.1	10.5	61	ASN	6.3	6.3	48	ILE	-	-	7.9	14.1
123	GLU	7.0	14.5	77	VAL	18.4	18.4	50	GLY	-	-	9.2	38.9
125	PRO	17.8	39.9	79	PHE	12.2	12.2	51	THR	-	-	31.8	76.0
126	GLU	28.8	59.3					52	SER	-	-	10.0	29.7
151	VAL	9.5	20.0					53	GLY	-	-	9.0	37.8
153	ASP	10.2	25.8					58	PHE	-	-	9.5	15.7
154	GLN	19.5	37.8					61	THR	-	-	6.9	16.4
159	TYR	14.6	23.5					65	HIS	-	-	6.7	12.1
161	VAL	16.5	34.6					87	ARG	31.3	43.1	16.7	23.0
191	LEU	6.2	11.0					88	PRO	14.2	31.9	-	-
192	ARG	40.0	55.0					90	SER	16.7	49.4	-	-
195	THR	12.3	29.4					92	LEU	20.1	35.6	17.7	31.3
196	CYS	6.3	15.1					93	TYR	-	-	7.3	11.8
201	SER	13.5	39.9					94	TYR	9.0	14.5	16.9	27.2
203	GLY	10.9	45.9					95	ASP	-	-	6.9	17.5
204	LEU	12.5	22.1					96	ASP	-	-	24.5	62.2
								97	GLY	-	-	10.1	42.7
								98	GLN	-	-	31.4	60.9
								99	ASN	11.8	28.8	10.3	25.1
								100	ILE	27.8	49.9	37.6	67.7
								101	ILE	11.7	21.0	7.2	13.0
								102	LYS	24.0	39.1	22.8	37.2
								103	LYS	-	-	8.2	13.4
								104	ASP	6.1	15.6	-	-
								105	ILE	-	-	10.9	19.6
								111	GLU	6.7	13.8	-	-

Table III.2. Amino acids of activin involved in binding to the receptor ActRIIB and follistatin. The serial numbers of the residues are given as identified using the protein complex structures of activin/follistatin and activin/ActRIIB.

Activin		ActRIIB	Follistatin
1	GLY		5,6,14,15
2	LEU		5,6,15
3	GLU		
5	ASP		31
8	VAL		17
10	ILE		17
25	TRP		
27	ASP		192
28	TRP		192
30	ILE	77,79	151,161,196
31	ALA	77,79	105,159,161
32	PRO	59,60,61,79	105,106,126,161
33	SER	61	126
36	HIS	60	125,126
47	HIS		15,16,17,46
48	ILE		14,15,16,17
50	GLY		15,16,44,46,47,48,50
51	THR		14,15,16,17,46,47,50
52	SER		5,6,14,15,16
53	GLY		6,14,15,47,50
58	PHE		47,50,51
61	THR		44,47
65	HIS		
87	ARG	30,33	123
88	PRO	30,57,58	
90	SER	38,55,56,57,58,59	106,126
92	LEU	38,79	105,106,153,159
93	TYR		153
94	TYR		52,153,192,201
95	ASP		48,51,52
96	ASP		48,52,203,204
97	GLY		203,204
98	GLN		191,192,195,201,203,204
99	ASN	13,16,17	153,154,192,195,201,203,204
100	ILE	13,51,52	151,152, 153,154,192,195,196,201,203
101	ILE	13,17,51,52	153,154,201
102	LYS	52,55	106,153,154
103	LYS		
104	ASP	55,56,57	
105	ILE		51
111	GLU		

Table III.3. Some selected top ranked representative follistatin ligands in the three hotspot regions (Reg) from the ZINC database using the LUDI docking runs. Absolute energy (ABS), percentage of fragment surface contact area with receptor (PCA), score from hydrogen bond term (HBS), lipophilic term (LPS), non hydrogen bonds (NHB), RMSD of compound fit and total Ludi3 score (L3S) are given. The percent solvent accessible contact area covered by each molecule on binding to follistatin residues is also given in parenthesis.

Reg	Molecules	ABS	PCA	HBS	LPS	NHB	RMSD	L3S	Follistatin AAs in contact with small molecules (in % area)
(i)	VZLF001	52.7	78	173	74	4	0.41	682	F52(8.5), K48(6.1)
	VZLF002	49.5	73	51	97	2	0.35	677	F52(11.1), K48(7.9)
	VZLF003	95.3	77	110	81	3	0.43	626	F52(11.1), K48(7.2)
	VZLF004	117.6	83	44	89	1	0.34	623	F52(6.7), K48(11.6)
	VZLF005	44.5	67	52	81	1	0.45	623	F52(15.2), K48(8.9)
	VZLF006	63.7	74	140	74	4	0.42	617	F52(7.6), K48(10.2)
	VZLF007	57.3	81	85	93	2	0.38	613	F52(7.5), K48(7.9)
	VZLF008	55.8	76	98	64	2	0.39	604	F52(12.8), K48(10.7)
	VZLF009	95.3	78	81	87	3	0.44	603	F52(11.1), K48(7.3)
	VZLF010	35.5	75	55	85	1	0.38	598	F52(18.3), K48(13.2)
	VZLF011	96.2	33	0	207	0	0.37	477	M50(17), F47(24.9), L46(6.8), N44(8.4), L16(10.2), V15(6.2)
	VZLF012	94.0	29	41	151	1	0.45	462	M50(17.2), F47(19.1), L46(7.3), L16(15.3), V15(9.2)
	VZLF013	128.8	26	52	153	1	0.35	404	F47(30.9), N44(8.9), Y17(6.3), L16(19), V15(25.1)
	VZLF014	82.0	43	110	110	2	0.35	403	K48(6.1), F47(35.7), L46(0.2), N44(25.4)
(ii)	VZLF015	146.4	26	42	151	1	0.37	408	Y159(14.8), D106(22.7), L105(10.3)
	VZLF016	80.6	26	49	135	1	0.4	399	Y159(10.1), E126(17.8), P125(1.1), D106(13.6), L105(21)
	VZLF017	72.0	33	38	195	1	0.39	393	Y159(17.4), E126(10), D106(28.7), L105(27.8)
	VZLF018	56.1	29	74	158	2	0.4	392	Y159(15.7), D106(24.5), L105(24.2)
	VZLF019	87.5	27	105	72	2	0.38	392	E126(23.9), P125(16.6)

Table III.3. (Continued…)

Reg	Molecules	ABS	PCA	HBS	LPS	NHB	RMSD	L3S	Follistatin AAs in contact with small molecules (in % area)
(iii)	VZLF020	5.9	34	91	103	2	0.4	480	L204(23.5), G203(14.5), R192(18.3), L191(9.1), A188(11.1)
	VZLF021	19.2	35	53	106	1	0.39	478	L204(8.6), G203(7.4), R192(20), L191(11), A188(16.6)
	VZLF022	89.9	28	102	122	3	0.43	478	V161(5.2), Q154(16.4), D153(19.5), V151(5.8)
	VZLF023	53.9	17	37	81	1	0.45	455	R192(21.3), V161(12.5), D153(10.6), V151(11.5)
	VZLF024	47.6	28	186	37	5	0.39	438	S201(13.8), T195(16.4),Q154(9.7), D153(25), V151(14.1)
	VZLF025	103.1	34	186	197	4	0.21	433	G203(13.4), S201(9.5), T195(14.8), R192(24.8), L191(9.8), A188(10.1)
	VZLF026	141.8	28	81	91	2	0.43	432	G203(9.7), S201(20.2), C196(12.2), T195(25.8), R192(23.5), V161(4.9), D153(1.9), V151(16.5)
	VZLF027	29.5	31	7	91	1	0.4	430	G203(15.2), S201(10.2), T195(13.5), R192(19.3), L191(8.5)
	VZLF028	32.2	22	71	62	2	0.32	427	L204(23.1), G203(17.5), R192(18.8), L191(5.3), A188(17.4)
	VZLF029	98.2	28	105	116	2	0.4	420	S201(20.1), T195(18.5), Q154(13.9), D153(25.1), V151(15)
	VZLF030	83.3	22	63	47	2	0.41	417	L204(15), G203(28.3), T195(6.3), R192(16.9), L191(9.9)
	VZLF031	26.3	31	154	79	4	0.34	417	G203(11.9), S201(4.4), T195(17.7), R192(28), L191(8.5)
	VZLF032	53.5	35	55	162	1	0.28	416	S201(19.2), T195(10.3), V161(11.9), Q154(21), D153(21.3), V151(14.8)

Table III.4. Some selected top ranking representative follistatin ligands in three hot spot regions (Reg) from the CAP database using LUDI docking runs. Absolute energy (ABS), percentage of fragment surface contact area with receptor (PCA), score from hydrogen bond term (HBS), lipophilic term (LPS), non hydrogen bonds (NHB), RMSD of compound fit and total Ludi3 Score (L3S) are given. The percent solvent accessible contact area covered by each molecule on binding to follistatin residues is also given in parenthesis.

Reg	Molecules	ABS	PCA	HBS	LPS	NHB	RMSD	L3S	Follistatin AAs in contact with small molecules (in % area)
(i)	VCLF001	36.2	79	55	83	1	0.43	722	F52(7.4), K48(8)
	VCLF002	68.5	75	140	81	3	0.39	711	F52(7.8), K48(6.8)
	VCLF003	56.6	84	55	87	1	0.44	687	F52(15.2), K48(8.7)
	VCLF004	103.9	79	36	124	2	0.4	650	F52(8.6), K48(7.5)
	VCLF005	225.9	71	35	43	1	0.45	643	F52(13.8), K48(16.3), N44(0.7)
	VCLF006	110.0	80	45	97	1	0.45	632	F52(13.4), K48(6.4)
	VCLF007	56.6	83	55	85	1	0.39	630	F52(5.5), K48(11.2)
	VCLF008	161.4	71	23	62	1	0.26	630	F52(4.7), K48(9.6)
	VCLF009	133.4	70	63	70	2	0.41	623	F52(1.9), K48(11.7)
	VCLF010	43.8	75	110	45	2	0.37	622	F52(13.3), K48(15.3)
	VCLF011	52.0	81	55	62	1	0.36	603	F52(5.5), K48(15.4), N44(1.4)
	VCLF012	63.9	75	115	66	3	0.37	600	F52(7.9), K48(8.8)
	VCLF013	98.0	77	103	62	2	0.34	600	F52(11.3), K48(9.6)
	VCLF014	98.4	76	81	81	2	0.43	597	F52(11.9), K48(10)
	VCLF015	115.7	21	0	139	0	0.39	464	M50(20.3), F47(37.4), L46(4.7), N44(9)
	VCLF016	80.8	35	0	180	0	0.34	450	M50(12.2), F47(33.9), L46(7.5), N44(12.9), L16(11.9)
	VCLF017	68.9	34	0	180	0	0.27	450	M50(12.7), F47(37.2), L46(6), N44(10.6), L16(11)
	VCLF018	52.1	32	49	164	1	0.43	428	M50(19.7), F47(27.4), L46(5.3), N44(8.5), L16(9.4), V15(7.9)
	VCLF019	115.7	30	0	193	0	0.37	408	I51(8.4), M50(30.6), F47(22.6), V15(11.5), Q14(12.1)

Table III.4. (Continued….)

Reg	Molecules	ABS	PCA	HBS	LPS	NHB	RMSD	L3S	Follistatin AAs in contact with small molecules (in % area)
(ii)	VCLF020	73.6	35	105	137	2	0.32	402	Y159(12.4), E126(11.7), D106(24.3), L105(14)
	VCLF021	52.2	23	55	114	1	0.43	384	Y159(18.3), D106(21.6), L105(24.9)
	VCLF022	135.1	25	55	114	1	0.4	384	Y159(19.7), E126(0.6), D106(28.9), L105(19.2)
	VCLF023	33.3	36	102	185	2	0.28	376	Y159(12.3), E126(18.8), D106(20.4), L105(27.1)
	VCLF024	60.7	23	54	106	1	0.35	375	Y159(20.8), D106(20.6), L105(24)
(iii)	VCLF025	141.8	28	81	91	2	0.43	432	V151(16.5), D153(1.9), V161(4.9), R192(23.5), T195(25.8), C196(12.2), S201(20.2), G203(9.7)
	VCLF026	36.2	40	128	103	6	0.39	422	A188(8.4), L191(11), R192(21.8), T195(17.6), S201(11.8), G203(23.2), L204(9.7)
	VCLF027	19.1	41	108	120	2	0.22	416	L191(7.1), R192(17.3), T195(10.8), S201(12.1), G203(18.3)
	VCLF028	19.7	36	41	95	1	0.27	412	A188(17.9), L191(8.4), R192(19.4), L204(8.4)
	VCLF029	68.5	22	56	58	2	0.32	408	L191(8.7), R192(19.8), T195(5.8), G203(22), L204(14.4)
	VCLF030	17.4	37	96	60	4	0.23	404	A188(6.7), L191(10.7), R192(21.4), T195(16.1), G203(13.1)
	VCLF031	19.0	20	68	45	2	0.4	403	L191(7), R192(19), T195(1.3), G203(22.3), L204(20.8)
	VCLF032	76.3	40	55	187	1	0.42	402	V151(19.4), D153(13.9), Q154(0.2), V161(15.2), R192(6.4), T195(22.4), C196(12.2), S201(15.1)
	VCLF033	49.0	34	163	189	5	0.39	402	A188(9), L191(9.6), R192(31.1), T195(17), S201(9.6), G203(13.5), L204(13.8)

Table III.5. Few selected top scoring representative follistatin ligands from the ZINC databases using the Glide-XP and the GOLD docking runs. Some selected and top ranked molecules using combined, Glide-XP (GlS), GOLD (GoS) and CombScore (CoS) scores in each of the hotspot region of follistatin are given. The solvent accessible contact area covered by each molecule on binding to follistatin residues is also given in parenthesis. The scores considered to select the representative molecules from the group are shown in **bold** letters.

Reg	Molecules	GoS	GlS	CoS	Follistatin AAs in contact with small molecules (in % contact area)
	VZGF001	40.2	-5.3	**68.5**	Q14(2.1), V15(10.9), L16(11.1), N44(11.1), L46(6), F47(28.1), M50(18.9)
	VZGF002	43.9	-4.3	**67.0**	Q14(8.9), V15(13), L16(18), N44(13.8), L46(6.8), F47(16.1), M50(21.7)
	VZGF003	36.6	-5.5	**66.1**	V15(1.5), L16(1.2), L46(0.9), F47(7.8), M50(11.6)
	VZGF004	34.5	-5.9	**66.1**	Q14(4.2), V15(13.4), L16(11.5), L46(6.2), F47(18.2), M50(35.8)
	VZGF005	40.6	-4.7	**65.9**	Q14(13.5), V15(14.5), L16(18.4), L46(6.8), F47(5.8), M50(35.3)
	VZGF006	36.6	-5.3	**65.1**	Q14(12.2), V15(12.4), L16(11.8), L46(6.4), F47(10.1), M50(35.4)
	VZGF007	34.7	-5.6	**64.8**	Q14(21), V15(20.2), L16(12.2), N44(0.1), L46(6.3), F47(12.6), M50(30.2)
	VZGF008	30.0	-6.4	**64.5**	Q14(14.8), V15(13.9), L16(12.9), N44(10.2), L46(6.6), F47(20.4), M50(33.1)
(i)	VZGF009	32.7	-5.9	**64.2**	Q14(15.4), V15(23.9), L16(16), N44(4.6), L46(6.8), F47(14.2), M50(23.8)
	VZGF010	32.6	-5.9	**64.2**	Q14(1.1), V15(19.5), L16(21), L46(5.8), F47(10.6), M50(17.1)
	VZGF011	17.0	**-6.6**	52.4	Q14(3.2), V15(12), L16(13), N44(0.6), L46(6.4), F47(19.3), M50(24.3)
	VZGF012	19.2	**-6.5**	54.2	Q14(3.4), V15(12), L16(12.9), N44(1.2), L46(6.4), F47(19.2), M50(24.4)
	VZGF013	13.9	**-6.5**	48.9	Q14(6.6), V15(11), L16(11.5), N44(6.1), L46(6), F47(22.4), M50(29.8)
	VZGF014	**39.9**	-4.4	63.4	Q14(13.4), V15(17.6), L16(16.4), L46(6.8), F47(11.3), M50(38.1)
	VZGF015	**38.0**	-4.1	60.3	Q14(13.7), V15(14.4), L16(12), N44(1.3), L46(6.2), F47(12.6), M50(38.4)
	VZGF016	**36.2**	-5.1	63.5	Q14(6.8), V15(13.3), L16(14.7), N44(0.5), L46(6.3), F47(6.8), M50(23.7)

Table III.5. (continued…)

Reg	Molecules	GoS	GlS	CoS	Follistatin AAs in contact with small molecules (in % contact area)
	VZGF017	56.1	-5.9	**101.0**	L105(25), D106(11.8), P125(2.9), E126(26), Y159(16.7)
	VZGF018	49.0	-6.5	**98.4**	L105(20.8), D106(20.8), E126(13.1), Y159(18.6)
	VZGF019	52.4	-5.8	**96.2**	L105(28.4), D106(17.1), E126(16), Y159(19.3)
	VZGF020	44.1	-6.6	**93.9**	L105(21.4), D106(16.8), E126(0.1), Y159(13.9)
	VZGF021	45.3	-6.2	**92.5**	L105(24), D106(17.8), E126(12.9), Y159(17.5)
	VZGF022	43.9	-6.4	**92.3**	L105(24.3), D106(19.9), E126(11.7), Y159(18.6)
(ii)	VZGF023	35.5	**-7.4**	91.4	L105(27.3), D106(16.8), E126(16.1), Y159(17)
	VZGF024	36.0	**-7.1**	89.6	L105(26.7), D106(19.4), E126(13.5), Y159(18.5)
	VZGF025	35.9	**-7.0**	89.0	L105(27.6), D106(16.2), E126(17.3), Y159(18.4)
	VZGF026	**47.8**	-5.3	88.1	L105(27.7), D106(19.2), E126(16.6), Y159(17.5)
	VZGF027	**47.8**	-4.2	79.5	L105(31.5), D106(0.7), Y159(12.8)
	VZGF028	**47.0**	-4.6	82.0	L105(26.9), D106(11.6), P125(2.4), E126(24.2), Y159(14.1)
	VZGF029	48.5	-6.1	**89.7**	V151(17.8), D153(11.9), V161(14.5), R192(22.6), T195(13.4), C196(20.3)
	VZGF030	56.3	-4.8	**89.1**	V151(19.7), D153(13.1), V161(21.4), R192(13.3), T195(17.6), C196(20.3), S201(1.9)
	VZGF031	45.0	-6.4	**88.3**	V151(13.3), D153(4.6), L191(10.8), R192(19.5), T195(30.3), C196(11.4), S201(17.3), G203(21.7), L204(10.9)
	VZGF032	38.1	-7.3	**87.6**	L191(10.6), R192(22.7), T195(11.5), S201(3.9), G203(19.6), L204(14.3)
	VZGF033	46.0	-6.1	**87.2**	V151(18.6), D153(17.1), V161(20.3), R192(10.2), T195(13.9), C196(20.3)
	VZGF034	51.1	-5.0	**85.3**	V151(18), D153(10.9), V161(15.6), R192(23.4), T195(12.8), C196(20.1
(iii)	VZGF035	47.2	-5.6	**85.3**	V151(14.7), D153(5.4), V161(3), R192(19), T195(19.6), C196(19.4), S201(13)
	VZGF036	31.2	**-6.9**	78.0	L191(10.8), R192(11.5), T195(11.6), S201(7.5), G203(20), L204(11.6)
	VZGF037	34.5	**-6.8**	80.8	L191(10.8), R192(10.8), T195(11.7), S201(7.5), G203(19.8), L204(16.2)
	VZGF038	26.0	**-6.8**	72.1	L191(10.8), R192(15.7), T195(12.5), S201(5.7), G203(22.7), L204(13.5)
	VZGF039	**48.9**	-4.0	75.8	L191(10.8), R192(19.1), T195(13.4), S201(5), G203(17.7), L204(14.9)
	VZGF040	**47.5**	-4.0	74.5	L191(10.8), R192(16.3), T195(10.7), S201(1.3), G203(19.6), L204(30.7)
	VZGF041	**47.4**	-4.2	75.8	V151(12.4), D153(12.2), Q154(0.1), L191(10.7), R192(19), T195(25), S201(25.4), G203(26.2)

Table III.6. Few selected top scoring representative follistatin ligands from the CAP database using the Glide-XP and the GOLD docking runs. Some selected and top ranked molecules using combined Glide-XP (GlS), GOLD (GoS) and CombScore (CoS) scores in each of the hotspot region (Reg) of follistatin are given. The solvent accessible contact area covered by each small molecule on binding to follistatin residues is also given in parenthesis. The scores considered during the selection of the representative molecules from the group are shown in **bold** letters.

Reg	Molecules	GoS	GlS	GoS	Follistatin AAs in contact with small molecules (in % area)
(i)	VCGF001	32.8	-5.7	**57.7**	Q14(0.1), V15(37.9), L16(20.7), Y17(9.4), L46(6), F47(11.5), M50(18.9)
	VCGF002	31.7	-5.6	**56.3**	V15(10.7), L16(17.8), N44(13), L46(6.4), F47(26.1), M50(12.4)
	VCGF003	29.0	-5.6	**53.6**	Q14(13.9), V15(13.8), L16(13.3), N44(0.4), L46(6.6), F47(17.2), M50(31.3)
	VCGF004	26.6	-5.4	**50.5**	V15(6.1), L16(14), N44(10.2), L46(6.3), F47(22.4), M50(14.2)
	VCGF005	26.6	-5.1	**49.0**	Q14(2.4), V15(19.4), L16(13.3), L46(6.4), F47(9.4), M50(22.3)
	VCGF006	25.0	-5.5	**49.0**	Q14(8.6), V15(14.6), L16(18.2), N44(12.2), L46(6.8), F47(24), M50(25.6)
	VCGF007	24.1	**-5.2**	47.0	Q14(12.1), V15(33.7), L16(14.6), L46(6.8), F47(7), M50(20.9)
	VCGF008	19.4	**-5.1**	41.9	Q14(3.4), V15(12.4), L16(20.7), L46(6.8), F47(8.8), M50(24.3)
	VCGF009	22.0	**-5.1**	44.4	Q14(1.9), V15(12.6), L16(12.3), N44(2.1), L46(6.3), F47(12.1), M50(20.5)
	VCGF010	**27.5**	-3.9	44.7	Q14(10.1), V15(28.4), L16(18.2), L46(6.5), M50(16.9)
	VCGF011	**26.9**	-4.7	47.4	Q14(0.2), V15(33.7), L16(20.9), Y17(0.6), N44(0.4), L46(6.3), F47(12.9), M50(19.1)
	VCGF012	**26.7**	-5.1	49.0	Q14(6.4), V15(24), L16(12.2), L46(6.2), F47(12.5), M50(21.1)
(ii)	VCGF013	41.5	-6.7	**93.2**	L105(22.2), D106(15.1), E126(14.1), Y159(17.7)
	VCGF014	49.8	-5.5	**92.1**	L105(21.5), D106(23), E126(11.9), Y159(19.8)
	VCGF015	41.1	-5.9	**86.2**	L105(21.8), D106(24.5), E126(11.4), Y159(20.6)
	VCGF016	36.4	-6.2	**84.2**	L105(10.5)
	VCGF017	42.2	-5.3	**83.1**	L105(20.1), D106(26.5), E126(12.2), Y159(19.9)
	VCGF018	35.7	-6.0	**82.1**	L105(10.6)
	VCGF019	39.5	-5.4	**81.3**	L105(15.4)
	VCGF020	29.3	**-6.0**	75.1	L105(10.7)
	VCGF021	27.5	**-5.9**	73.2	L105(15.6)
	VCGF022	35.6	**-5.8**	80.0	L105(16.3)
	VCGF023	31.1	**-5.7**	74.5	L105(10.5)
	VCGF024	36.0	**-5.5**	78.1	L105(20.8), D106(25.7), E126(11.2), Y159(20.9)
	VCGF025	**45.6**	-3.5	72.6	L105(21.3)
	VCGF026	**40.3**	-4.3	73.6	L105(13.8)
	VCGF027	**40.0**	-3.7	68.8	L105(17.9)
	VCGF028	**40.0**	-4.1	71.1	L105(18.7)

Table III.6. (Continued…)

Reg	Molecule	GoS	GlS	GoS	Follistatin AAs in contact with small molecules (in % area)
(iii)	VCGF029	36.9	-6.9	**78.8**	L191(10.8), R192(11.7), T195(12.9), S201(7.4), G203(18.9), L204(20.9)
	VCGF030	46.0	-5.2	**77.7**	V151(19.7), D153(20.9), Q154(2.3), V161(18.2), R192(4.7), T195(16), C196(15.9), S201(7.6)
	VCGF031	44.7	-5.1	**75.7**	L191(10.8), R192(10.5), T195(11.7), S201(9), G203(29.1), L204(27.2)
	VCGF032	35.3	-6.6	**75.1**	V151(16.8), L191(10.8), R192(18.1), T195(10.1), S201(5.4), G203(20.4), L204(12.4)
	VCGF033	37.0	-6.2	**74.4**	L191(10.8), R192(8.9), T195(10.2), S201(13), G203(23.4), L204(13.9)
	VCGF034	36.3	-6.1	**73.3**	L191(10.8), R192(17.6), T195(21.9), S201(16.3), G203(21.4), L204(14.5)
	VCGF035	33.4	-6.5	**72.9**	L191(10.8), R192(16.9), T195(21.9), S201(17.2), G203(23.9), L204(14.3)
	VCGF036	33.3	-6.2	**70.9**	L191(10.8), R192(19.1), T195(14.8), S201(6.5), G203(19.3), L204(14.4)
	VCGF037	31.6	-6.5	**70.9**	L191(10.8), R192(16), T195(13.5), S201(11), G203(19.6), L204(16.1)
	VCGF038	40.8	-4.9	**70.2**	V151(19.3), D153(10), V161(18.2), R192(16.3), T195(18.4), C196(20.3), S201(5.5)
	VCGF039	32.0	**-6.3**	70.2	L191(10.8), R192(10), T195(9.3), S201(13.7), G203(27.1), L204(17.1)
	VCGF040	30.5	**-6.2**	68.0	L191(10.8), R192(11.6), T195(10.7), S201(12.5), G203(27.2), L204(11.5)
	VCGF041	32.2	**-6.2**	69.4	L191(10.8), R192(14.6), T195(13.8), S201(10.3), G203(19.1), L204(18.8)
	VCGF042	29.6	**-6.1**	66.6	R192(17.3), T195(22.1), S201(17.2), G203(22.1), L204(13.1)
	VCGF043	31.5	**-6.1**	68.2	L191(10.8), R192(16.1), T195(13.7), S201(12.1), G203(20.1), L204(15)
	VCGF044	**36.9**	-4.3	62.5	V151(18.2), D153(4.4), V161(20.3), R192(14.6), T195(20.4), C196(20.3), S201(15.1)
	VCGF045	**36.8**	-4.8	65.7	L191(10.8), R192(19.7), T195(9.6), S201(5.5), G203(22.1), L204(15.4)
	VCGF046	**36.6**	-4.8	65.4	L191(10.8), R192(16.9), T195(9.8), G203(20.4), L204(17.4)
	VCGF047	**36.1**	-4.4	62.8	L191(10.8), R192(19.1), T195(11), S201(6.5), G203(22.8), L204(19.2)
	VCGF048	**36.0**	-5.7	70.1	V151(19.6), D153(11.8), V161(17.5), R192(12.7), T195(20.3), C196(20.3), S201(5.5)

Table III.7. Binding energy (BE) values calculated using Molecular Mechanics Poisson-Boltzmann Surface Area (MM-PBSA) method for the top scoring representative follistatin ligands from the ZINC and the CAP databases given in Table III.5 and III.6.

ZINC Molecules from Table III.5				**CAP Molecules from Table III.6**			
Molecule	**BE**	**Molecule**	**BE**	**Molecule**	**BE**	**Molecule**	**BE**
Region (i)		Region (iii)		Region (i)		Region (iii)	
VZGF001	-12.55	VZGF029	-18.33	VCGF001	-5.21	VCGF029	-6.47
VZGF002	32.71	VZGF030	6.09	VCGF002	-10.87	VCGF030	-0.89
VZGF003	-144.00	VZGF031	-19.92	VCGF003	-25.94	VCGF031	-11.69
VZGF004	-26.74	VZGF032	-20.42	VCGF004	-1.19	VCGF032	-12.78
VZGF005	-4.96	VZGF033	-4.14	VCGF005	-15.77	VCGF033	-1.29
VZGF006	-26.86	VZGF034	-15.50	VCGF006	-26.24	VCGF034	0.00
VZGF007	-20.46	VZGF035	3.66	VCGF007	-16.86	VCGF035	-0.97
VZGF008	-34.65	VZGF036	-6.94	VCGF008	-3.15	VCGF036	-13.21
VZGF009	-20.26	VZGF037	-2.77	VCGF009	-25.94	VCGF037	-21.71
VZGF010	-1.74	VZGF038	-22.56	VCGF010	10.34	VCGF038	-6.04
VZGF011	-32.57	VZGF039	5.80	VCGF011	-18.34	VCGF039	88.96
VZGF012	-30.30	VZGF040	-0.01	VCGF012	-7.10	VCGF040	-13.03
Region (ii)		VZGF041	-1.18	VCGF013	-13.86	VCGF041	-20.34
VZGF013	-16.46			VCGF014	-5.26	VCGF042	-0.38
VZGF014	-23.46			VCGF015	-0.58	VCGF043	-21.88
VZGF015	-4.28			VCGF016	-11.97	VCGF044	36.42
VZGF016	-10.34			Region (ii)		VCGF045	-15.73
VZGF017	-21.81			VCGF017	-11.81	VCGF046	-9.67
VZGF018	-9.06			VCGF018	-0.07	VCGF047	-4.35
VZGF019	-7.70			VCGF019	-17.34	VCGF048	-8.36
VZGF020	-12.02			VCGF020	-18.42		
VZGF021	-8.62			VCGF021	-12.65		
VZGF022	1.38			VCGF022	-15.04		
VZGF023	-11.77			VCGF023	-6.14		
VZGF024	-4.17			VCGF024	3.13		
VZGF025	-7.12			VCGF025	-8.02		
VZGF026	-7.39			VCGF026	-16.25		
VZGF027	-12.37			VCGF027	-12.11		
VZGF028	-20.46			VCGF028	-1.11		

Table IV.1. The contact regions of amino acids of CV2, BMPRIA and BMP2**.** Surface accessible contact areas (Sum), in percent (% SA) of residues involved in binding with BMP-2 are given. Only the residues that show a loss of greater than10 % SA are given.

CV2 (3bk3)				BMPRIA (2h62)				BMP-2 contact with					
Pos	**AA**	**Sum**	**%SA**	**Pos**	**AA**	**Sum**	**%SA**	**Pos**	**AA**	**Sum**	**%SA**	**Sum**	**%SA**
										Receptors		**With CV2**	
-1	TRP	25.3	33.6	42	GLY	10.1	42.5	10	LEU	29.2	51.7	-	-
1	LEU	28.6	50.6	43	HIS	22.1	40.2	25	ASP	9.0	22.8	-	-
2	ILE	49.2	88.5	45	PRO	13.1	29.4	26	VAL	7.3	15.4	-	-
3	THR	9.4	22.4	46	ASP	4.0	10.1	27	GLY	4.7	19.5	-	-
4	GLY	14.7	61.9	67	ASP	11.8	30.0	28	TRP	14.8	19.6	11.0	14.6
5	THR	13.8	32.9	72	THR	4.6	11.0	30	ASP	4.5	11.5	-	-
6	GLU	15.2	31.4	77	CYS	9.7	23.1	31	TRP	19.5	25.9	22.0	29.2
7	ALA	17.9	53.8	79	LYS	25.2	41.0	33	VAL	14.8	31	10.3	21.6
18	ILE	10.3	18.6	81	GLU	7.9	16.2	34	ALA	12.2	36.7	10.8	32.6
20	ASN	8.9	21.6	82	GLY	10.9	46.0	35	PRO	8.8	19.6	10.2	23.0
21	ILE	28.1	50.6	84	ASP	20.6	52.4	39	HIS	18.5	33.7	-	-
27	ILE	13.1	23.5	85	PHE	51.3	84.4	49	PHE	29.7	48.9	-	-
36	ALA	6.4	19.2	86	GLN	28.6	55.4	50	PRO	30.7	69	-	-
38	CYS	16.6	39.6	88	LYS	27.4	44.6	51	LEU	6.2	11.1	7.3	12.9
40	GLN	14.1	27.3	89	ASP	9.8	24.7	52	ALA	11.5	34.5	-	-
				90	SER	17.6	52.1	53	ASP	36.0	91.4	-	-
				91	PRO	5.1	11.3	54	HIS	27.4	49.8	-	-
				92	LYS	26.1	42.6	57	SER	8.2	24.3	-	-
				93	ALA	3.6	10.9	59	ASN	7.2	17.6	9.9	24.2
				94	GLN	23.6	45.8	62	ILE	20.2	36.3	10.9	19.5
				97	ARG	15.8	21.8	66	LEU	13.5	23.9	-	-
				BMPR-II (2hlr)				69	SER	18.08	53.6	-	18.1
				Pos	**AA**	**Sum**	**%SA**	70	VAL	19.3	40.4	-	-
				40	TYR	23.3	37.5	72	SER	6.2	18.4	-	-
				67	TYR	6.5	10.4	76	LYS	12.3	19.9	-	-
				69	LEU	11.7	20.7	85	SER	7.3	21.6	-	-
				81	LYS	15.2	24.8	86	ALA	7.8	23.6	-	-
				84	CYS	8.8	21	87	ILE	9.7	17.5	-	-
				85	TRP	20.8	27.7	88	SER	15.8	46.8	9.1	27.0
				86	SER	15.3	45.4	90	LEU	17.2	30.5	18.2	32.2
				87	HIS	18.9	34.3	96	GLU	9.1	18.8	-	-
				88	ILE	19.1	34.5	97	LYS	13.8	22.5	-	-
				89	GLY	12.2	51.4	98	VAL	21.5	45.1	18.9	39.6
				90	ASP	15.5	39.5	99	VAL	-	-	8.9	18.7
				93	GLU	7.9	16.3	100	LYS	23.1	37.6	34.4	60.8
				113	TYR	16.4	26.3	101	LYS	11.5	18.7	15.8	25.8
				115	PHE	13.0	21.4	102	ASP	6.3	16	16.7	40.9
								103	TYR	10.1	16.2	11.5	18.5
								107	VAL	4.8	10.1	-	-

Table IV.2. Amino acids in BMP-2 that are involved in binding to receptors BMPR-IA, BMPR-II, and CV2 calculated using the modeled protein complexes of BMP-2/BMPR-IA/BMPR-II and BMP-2/CV2.

BMP-2		**BMPR-IA**	**BMPR-II**	**CV2**
10	LEU	42-43,67,72		
25	ASP	90-92		
26	VAL	85,88-93		1
27	GLY	85,88-92		1
28	TRP	85,88-91		-1 - 2
30	ASP	88		-1 - 2
31	TRP	84-85,88		-1
33	VAL		69,113,115	20, 21, 27, 38, 40
34	ALA		67,69,85,93,115	18, 20, 21, 27, 38, 40
35	PRO		85,87,93,115	18, 20, 21, 27
36	PRO		93	20, 21
37	GLY			20, 21
38	TYR			20, 21
39	HIS		87-90,93	20
49	PHE	43,86,89,97		
50	PRO	43,77,82,85-86,89		1
51	LEU	43,77,82,85-86		1, 2
52	ALA	42-43,45,77,79,82,86		1
53	ASP	43,45,77,79,81-82		-1, 1
54	HIS	43,45-46,77,79		
57	SER	79,81-82		1 - 3
58	THR			1 - 4
59	ASN	79,81-82,84-85		-1 - 4
60	HIS			1 - 3

Table IV.2. (Continued…)

	BMP-2	BMPR-IA	BMPR-II	CV2
61	ALA			1 - 3
62	ILE	82,84-86,89		-1 - 3
63	VAL			1 – 3
66	LEU	85-86,89-90,92-94		1, 2
69	SER	89-94		
70	VAL	89-94		
72	SER	93-94		
85	SER		86-89	
86	ALA		85-89	3 - 5, 7
87	ILE		85-90	3 - 7, 18, 20, 21
88	SER		67,84-88	3 - 7,18, 20, 21, 36
89	MET			3 - 7, 18, 20, 21, 27, 36, 38
90	LEU		67,69,81,85,115	3 - 7, 18, 20, 21, 27, 36, 38, 40
91	MET			-1, 2 - 7, 21, 38, 40
93	ASP			-1, 4, 6
96	GLU		40	40
97	LYS		40,81	38, 40
98	VAL		40,69,81	6, 27, 36, 38, 40
99	VAL		40,81	4 - 7, 21, 27, 36, 38, 40
100	LEU		40,81,84-85	4 - 7, 18, 21, 27, 36, 38
101	LYS	81	84-86	2 - 7, 18, 21, 36, 38
102	ASN	81	84-87	2 - 7, 18, 21, 36
103	TYR	81-82	86	1 - 7
104	GLN	81-82	86	2 - 7
105	ASP			2 - 4
106	MET			2 - 4
107	VAL		87-90	

Table IV.3. Some selected high Ludi3 score CV2 ligands from the CAP database using the LUDI docking runs. Absolute energy (ABS), percentage of fragment surface contact area with receptor (PCA), score from hydrogen bond term (HBS), lipophilic term (LPS), non hydrogen bonds (NHB), RMSD of compound fit and total Ludi3 Score (L3S) are given.

Molecule	ABS	PCA	HBS	LPS	NHB	RMSD	L3S	CV2 residues that bind to BMP-2 (% residue surface in contact with small molecule)
VCLC001	51.3	20	102	43	3	0.36	525	W-1(21.9), L1(9.8), I2(15.5)
VCLC002	43.8	30	48	91	1	0.35	448	W-1(21), L1(10.5), I2(24.7)
VCLC003	40.4	22	55	49	1	0.43	429	W-1(34.8), L1(4), I2(17.2)
VCLC004	37.2	23	110	99	2	0.22	424	W-1(21.9), L1(2.5), I2(20.5), G4(56), T5(1), E6(8.5)
VCLC005	44.3	13	55	60	2	0.44	424	W-1(28.1), L1(0.1), I2(25.5)
VCLC006	128.2	26	55	151	1	0.40	421	W-1(19.3), L1(4.8), I2(32.3), G4(40.2), T5(8.6), E6(6.3), A7(5.5)
VCLC007	16.2	23	55	99	1	0.38	408	W-1(16.3), L1(4.2), I2(12.6)
VCLC008	41.2	27	55	72	1	0.42	397	W-1(16.8), L1(7.9), I2(13.7)
VCLC009	41.5	25	55	66	1	0.38	391	W-1(20), L1(5.7), I2(13.2)
VCLC010	52.0	16	54	60	1	0.44	384	W-1(25.3), L1(1.5), I2(22.6)
VCLC011	41.2	17	51	60	1	0.41	381	W-1(26.9), L1(5), I2(16.4)
VCLC012	20.8	25	55	91	1	0.43	361	W-1(16.4), L1(9), I2(26.3)
VCLC013	151.7	32	113	85	5	0.64	358	W-1(13.9), I2(15.7), T3(6.7), G4(71.4), T5(6.3), E6(12.8)
VCLC014	14.9	23	55	97	1	0.41	351	W-1(19.2), L1(2.4), I2(14.2)
VCLC015	31.3	24	17	58	1	0.41	345	W-1(14.3), L1(9.2), I2(25.4)
VCLC016	154.9	21	55	72	1	0.41	342	W-1(19.8), L1(11), I2(20.4)
VCLC017	57.6	16	108	68	2	0.25	336	W-1(20.9), I2(12.7), G4(39.1), T5(8), E6(15.9), A7(5.4)
VCLC018	140.0	37	79	97	3	0.39	336	W-1(17.9), I2(4), T3(37.3), G4(44.1), T5(11.6)
VCLC019	53.0	24	122	54	4	0.42	336	W-1(13.2), I2(11.6), T3(9.3), G4(49.4), E6(2.3)
VCLC020	7.5	24	158	16	4	0.37	334	W-1(13.9), L1(1.6), I2(14.8), G4(0.3)
VCLC021	59.0	22	49	85	1	0.41	333	W-1(20.9), I2(20.7), G4(22.6)
VCLC022	128.7	36	79	91	3	0.39	330	W-1(17.9), I2(3.9), T3(35.4), G4(44.1), T5(6.7)
VCLC023	43.3	26	122	47	4	0.42	329	W-1(13.2), I2(11.5), T3(9.3), G4(45.1)
VCLC024	43.8	20	55	74	1	0.30	328	W-1(18.7), L1(0.9), I2(25.3)
VCLC025	38.8	25	110	47	3	0.34	317	W-1(10.9), L1(9.1), I2(32.9)
VCLC026	41.5	25	55	99	1	0.33	314	W-1(13.3), L1(4.7), I2(21.7)
VCLC027	47.6	24	55	114	1	0.31	313	W-1(17.2), L1(4.9), I2(23.2)
VCLC028	20.4	25	90	118	2	0.33	313	W-1(13.7), I2(13.2), G4(54.9), T5(2.6), E6(14.1)
VCLC029	26.2	38	55	112	1	0.28	311	W-1(15.8), L1(6.4), I2(17.5)
VCLC030	5.8	39	110	39	3	0.36	309	W-1(11), L1(9.7), I2(12.5)

Table IV.4. Some selected high Ludi3 score CV2 ligands from the ZINC database using the LUDI docking runs. Percentage of the fragment surface contact area with receptor (PCA), score from hydrogen bond term (HBS), lipophilic term (LPS), non hydrogen bonds (NHB), RMSD of compound fit and total Ludi3 Score (L3S) are given.

Molecules	PCA	HBS	LPS	HB	RMSD	L3S	CV2 residues in contact with small molecule (percentage of contact area covered)
VZLC001	17	55	66	1	0.28	556	W-1(21.7), L1(1.4), I2(26.1)
VZLC002	16	55	56	1	0.42	546	W-1(30), L1(1.9), I2(13.1)
VZLC003	17	35	74	1	0.44	544	W-1(26.6), L1(6.6), I2(16.5), G4(26.4)
VZLC004	21	56	43	2	0.43	541	W-1(39.3, L1(3.3), I2(15.3)
VZLC005	26	55	47	1	0.38	537	W-1(26.4), L1(12.2), I2(18.1)
VZLC006	37	54	101	2	0.27	535	W-1(13.2), L1(10.6), I2(27.6), G4(3.2)
VZLC007	12	55	41	1	0.41	531	W-1(32.9), L1(7.1), I2(17.8)
VZLC008	26	46	45	1	0.34	526	W-1(27.6), L1(11.8), I2(17.7)
VZLC009	20	81	106	3	0.42	519	W-1(27), L1(9.2), I2(14.4), G4(41), T5(2.4), E6(9.4)
VZLC010	19	55	76	1	0.4	511	W-1(29.2), L1(8.3), I2(19.2)
VZLC011	33	54	74	1	0.45	508	W-1(17.9), L1(9.3), I2(18.6)
VZLC012	31	34	124	1	0.41	506	W-1(25.4), L1(5.8), I2(24.7)
VZLC013	21	13	101	1	0.41	501	W-1(37.4), L1(4.8), I2(20.7), T3(0.3), G4(37)
VZLC014	19	58	58	2	0.41	496	W-1(31.7), L1(4.7), I2(15.8)
VZLC015	14	55	58	1	0.21	493	W-1(28.5), L1(3.2), I2(25.8)
VZLC016	13	0	33	0	0.39	491	W-1(46.9), I2(1.9)
VZLC017	20	51	58	2	0.42	489	W-1(32), L1(4.1), I2(16.6)
VZLC018	16	55	47	2	0.43	489	W-1(36.2), L1(2.7), I2(15)
VZLC019	30	55	51	1	0.35	486	W-1(20.3), L1(14), I2(18.7)
VZLC020	18	0	60	0	0.41	486	W-1(36.3), L1(13.7), I2(22.4)
VZLC021	22	24	128	1	0.42	484	W-1(25.3), L1(10.5), I2(30.8), T3(0.5), G4(20.5)
VZLC022	17	55	62	1	0.41	481	W-1(32.3), L1(1.4), I2(23.3), G4(2.5)
VZLC023	24	32	81	1	0.41	477	W-1(28), L1(5), I2(17.4)
VZLC024	29	54	97	1	0.37	476	W-1(22.8), L1(8.2), I2(28.2), G4(14.9)
VZLC025	22	47	101	1	0.32	473	W-1(26.8), L1(8.2), I2(21.6), G4(23.1)
VZLC026	11	55	31	1	0.44	466	W-1(32.4), L1(7.9), I2(5.1)
VZLC027	18	55	62	1	0.37	465	W-1(25.2), L1(7.8), I2(21.8)
VZLC028	24	93	91	2	0.38	374	T5(6.2), E6(24.9), A7(22.5), A36(17.9), C38(2.7)
VZLC029	25	100	145	2	0.32	357	W-1(11.9), L1(6.1), I2(25.1), G4(31.3), T5(6.4), E6(10.4), A7(18.8), A36(11.1), C38(2)
VZLC030	18	55	85	1	0.42	394	W-1(31.4), I2(2.8), G4(32.6), T5(11.7), E6(7.9), A7(30.1), I18(10.3), I21(5.1)

Table IV.5. Few selected top scoring representative CV2 ligands from the ZINC database using the Glide-XP and the GOLD docking runs. Some selected and top ranked molecules using combined Glide-XP (GlS) and GOLD (GoS) and CombScore (CoS) scores in the hotspot region of CV2 are given. The solvent accessible contact area covered by each small molecule on binding to CV2 residues is also given in parenthesis. The scores considered during the selection of the representative molecules from the group are shown in **bold** letters.

Molecules	**GoS**	**GlS**	**CoS**	**CV2 residues in contact with small molecule (percentage of contact area covered)**
VZGC001	47.9	-6.4	**117.8**	E6(21.1), A7(33), I18(12.3), I21(10.2), A36(20.5), C38(22.4)
VZGC002	50.9	-6.1	**117.0**	E6(16.3), A7(33.8), I18(18.2), I21(12.8), A36(20.8), C38(16.3)
VZGC003	54.5	-5.7	**116.1**	G4(5.3), T5(8), E6(20.8), A7(33.3), I18(12.7), I21(17.9), A36(20.4), C38(27.3)
VZGC004	45.3	-6.5	**115.7**	T5(2.5), E6(18.5), A7(43.9), I18(20.5), I21(11.3), A36(20.8), C38(23.3)
VZGC005	50.3	-6.0	**115.2**	T5(1), E6(20.2), A7(35.1), I18(12.6), I21(8.5), A36(20.7), C38(32.1)
VZGC006	51.7	-5.8	**115.1**	E6(15.7), A7(41), I18(21), I21(10), A36(20.8), C38(16)
VZGC007	41.4	-6.8	**114.9**	T5(3.8), E6(19.1), A7(42.2), I18(16.2), I21(20.6), A36(20.7), C38(16.1)
VZGC008	47.3	-6.2	**114.9**	T5(2.5), E6(20.5), A7(34.1), I18(12.9), I21(8.3), A36(20.5), C38(22.9)
VZGC009	49.8	-5.8	**113.1**	E6(17.5), A7(30.9), I18(13.5), I21(17.4), A36(20.7), C38(16.1)
VZGC010	53.3	-5.5	**113.0**	E6(18.6), A7(35.8), I18(13.9), I21(9.6), A36(20.5), C38(39.5)
VZGC011	29.8	**-6.3**	98.5	T5(7.9), E6(26), A7(24), I18(3.3), I21(2.2), A36(19.1), C38(34.8)
VZGC012	31.4	**-6.3**	100.0	T5(6.9), E6(18.7), A7(30.2), I18(10.9), I21(18.5), A36(19.3), C38(26.9)
VZGC013	40.3	**-6.2**	108.1	E6(14.4), A7(35.9), I18(19.6), I21(16.1), A36(20.6), C38(26)
VZGC014	37.9	**-6.0**	103.5	E6(14.7), A7(41.8), I18(26), I21(11.6), A36(20.9), C38(25)
VZGC015	41.6	**-6.0**	106.9	E6(17.8), A7(34.3), I18(12.6), I21(22.5), I27(14.4), A36(20.4), C38(29.1), Q40(9)
VZGC016	32.1	**-6.0**	97.2	T5(0.2), E6(20.6), A7(34.5), I18(15.6), I21(15), A36(20.4), C38(25.5)
VZGC017	32.1	**-5.9**	96.8	T5(0.8), E6(20), A7(36.4), I18(18.3), I21(8.2), A36(19.9), C38(25.4)
VZGC018	38.9	**-5.9**	103.4	E6(15), A7(39.6), I18(17.6), I21(11.7), A36(20.6), C38(28.2)
VZGC019	23.0	**-5.9**	87.0	T5(0.6), E6(15.5), A7(31.4), I18(10.3), I21(14.8), A36(20.1), C38(16.4)
VZGC020	46.1	**-5.8**	109.1	E6(15.3), A7(42.4), I18(23), I21(11.3), A36(21), C38(26.3)
VZGC021	**56.1**	-4.3	102.3	E6(15.2), A7(41.7), I18(27.1), I21(11.8), A36(20.9), C38(23)
VZGC022	**55.5**	-1.8	75.2	E6(15.2), A7(32.4), I18(17.4), I21(22.5), I27(14.2), A36(20.4), C38(32.3), Q40(7.5)
VZGC023	**55.4**	-2.0	76.8	E6(8.2), A7(26.9), I18(15.9), I21(21.8), I27(9.6), A36(19), C38(43.7), Q40(11.5)
VZGC024	**54.0**	-3.4	91.0	E6(17.4), A7(40.5), I18(18.7), I21(19.7), I27(5.5), A36(20.7), C38(34.7)
VZGC025	**53.1**	-2.4	79.0	T5(1.6), E6(20.8), A7(34.7), I18(13.7), I21(9.8), A36(20.8), C38(15.6)
VZGC026	**52.7**	-4.3	99.8	T5(0.5), E6(24.1), A7(34.7), I18(20.4), I21(14.6), A36(20.5), C38(12.8)
VZGC027	**52.6**	-5.0	107.1	E6(14.8), A7(40.5), I18(23), I21(7.7), A36(21), C38(18.2)
VZGC028	**52.5**	-5.5	112.5	E6(17.6), A7(31.7), I18(13.7), I21(15.9), A36(20.7), C38(18.9)
VZGC029	**52.3**	-5.3	110.0	T5(1.5), E6(18.4), A7(42.1), I18(18.4), I21(10.5), A36(20.4), C38(20)
VZGC030	**52.1**	-2.4	78.4	T5(0.5), E6(14.3), A7(35.9), I18(20.4), I21(12.8), A36(20.6), C38(17.1)

Table IV.6. Few selected top scoring representative CV2 ligands from the CAP database using the Glide-XP and the GOLD docking runs. Some selected and top ranked molecules using combined Glide-XP (GlS), GOLD (GoS) and CombScore (CoS) scores in the hotspot region of CV2 are given. The solvent accessible contact area covered by each small molecule on binding to CV2 residues is also given in parenthesis. The scores considered during the selection of the representative molecules from the group are shown in **bold** letters.

Molecules	GoS	GlS	CoS	CV2 residues in contact with small molecule (percentage of contact area covered)
VCGC001	35.3	-5.4	**89.2**	T5(1.8), E6(11.4), A7(37.5), I18(12.3), I21(10.3), A36(20.4), C38(26.2)
VCGC002	34.0	-5.5	**88.9**	T5(4.3), E6(14.1), A7(37.5), I18(11.9), I21(8.9), A36(20.4), C38(26.1)
VCGC003	43.4	-4.4	**87.5**	A7(28), I18(14.8), I21(22.6), I27(11.7), A36(5.9), C38(31.1), Q40(16.1)
VCGC004	42.1	-4.4	**86.2**	E6(11.8), A7(17.6), I21(21), I27(16.4), A36(18.3), C38(33.8), Q40(6)
VCGC005	38.3	-4.4	**82.8**	A7(21.8), I18(12.9), I21(25), I27(12), A36(5.9), C38(30.7), Q40(14.4)
VCGC006	39.4	-4.3	**82.7**	E6(9.9), A7(34.3), I18(17.4), I21(23.4), I27(11), A36(20.7), C38(36.2), Q40(10.4)
VCGC007	42.5	-4.0	**82.3**	E6(17.6), A7(35.5), I18(12.7), I21(25.3), I27(12.8), A36(20.6), C38(35.3), Q40(6.9)
VCGC008	37.9	-4.4	**82.2**	E6(10.8), A7(42.1), I18(19.2), I21(19.2), I27(12.9), A36(16.6), C38(33.2), Q40(7.7)
VCGC009	33.7	-4.8	**82.2**	E6(10.4), A7(31.9), I18(10.2), I21(13.4), A36(18), C38(22.5)
VCGC010	42.4	-3.9	**82.0**	G4(22.4), T5(7.7), E6(22.8), A7(42.6), I18(19.3), I21(20.1), I27(9), A36(11.1), C38(33.4), Q40(8.6)
VCGC011	24.8	**-5.3**	77.8	T5(2.9), E6(16.1), A7(33.2), I18(11.8), I21(8.1), A36(18.8), C38(35.2)
VCGC012	27.4	**-5.0**	77.3	E6(10.3), A7(36.9), I18(14.5), I21(17.3), A36(18.9), C38(28.6)
VCGC013	28.9	**-4.9**	78.7	E6(12.3), A7(42.9), I18(20.5), I21(19.6), I27(2.9), A36(19.9), C38(27.3)
VCGC014	18.9	**-4.9**	68.3	E6(10.4), A7(33.2), I18(12.6), I21(10.4), A36(18.8), C38(22.8)
VCGC015	27.4	**-4.9**	76.7	A7(38.1), I18(28.4), I21(30.4), A36(7.1), C38(19.7)
VCGC016	24.6	**-4.9**	73.7	E6(13.8), A7(37.5), I18(18), I21(23.8), A36(19.4), C38(23.6)
VCGC017	33.0	**-4.8**	81.8	E6(15.2), A7(37.4), I18(21.3), I21(18.1), A36(19.2), C38(12.7)
VCGC018	28.8	**-4.9**	77.6	E6(20.1), A7(30.5), I18(10.3), I21(9.1), A36(20.5), C38(27.3)
VCGC019	27.6	**-4.8**	76.5	E6(19.9), A7(32.6), I18(12.9), I21(16.4), A36(20.6), C38(26.3)
VCGC020	23.9	**-4.8**	72.5	T5(7.2), E6(14.7), A7(31.7), I18(8.1), I21(13.6), A36(20.2), C38(23.1)
VCGC021	**44.8**	-2.9	74.3	A7(29.4), I18(16.2), I21(25.5), I27(19.4), A36(5.1), C38(24.7), Q40(16.7)
VCGC022	**44.4**	-2.2	66.0	A7(42.5), I18(30.1), I21(32.1), I27(14.8), A36(13.7), C38(31), Q40(10.3)
VCGC023	**44.2**	-2.6	70.2	A7(37.4), I18(29.5), I21(25.4), I27(9.8), A36(9.1), C38(37.5), Q40(8.8)
VCGC024	**44.0**	-2.2	65.8	A7(30.1), I18(16.3), I21(27.9), I27(11.9), A36(5.3), C38(30.9), Q40(12.9)
VCGC025	**43.9**	-3.6	79.8	T5(8.3), E6(23.5), A7(34.9), I18(12.9), I21(9.9), A36(20.4), C38(12)
VCGC026	**43.7**	-2.3	67.0	A7(43.2), I18(29.7), I21(29.6), I27(16.5), A36(13.6), C38(30), Q40(9.6)
VCGC027	**43.3**	-2.3	66.9	T5(6.9), E6(29), A7(37), I18(15.2), I21(7.4), A36(19.2), C38(31.5)
VCGC028	**43.0**	-3.1	74.4	E6(4.9), A7(32), I18(21.7), I21(22.7), I27(10.8), A36(13.3), C38(35.6), Q40(16.9)
VCGC029	**42.9**	-2.3	66.3	A7(34.7), I18(20.6), I21(29.5), I27(7.5), A36(6), C38(28.1), Q40(10.3)
VCGC030	**42.6**	-3.7	79.4	A7(27.3), I18(16.5), I21(25.4), I27(12.4), A36(5.7), C38(31.3), Q40(15)

Table IV.7. Binding energy values calculated using Molecular Mechanics Poisson-Boltzmann Surface Area (MM-PBSA) method for the top scoring representative CV2 ligands from the ZINC and the CAP databases given in Table IV.5 and IV.6.

ZINC Molecules in Table IV.5	MM-PBSA Binding Energy	CAP Molecules in Table IV.6	MM-PBSA Binding Energy
VZGC001	-3.81	VCGC001	-13.87
VZGC002	-2.98	VCGC002	-8.02
VZGC003	-2.43	VCGC003	-34.30
VZGC004	-9.31	VCGC004	-18.77
VZGC005	-8.85	VCGC005	-32.68
VZGC006	-11.44	VCGC006	-32.37
VZGC007	-11.44	VCGC007	-15.79
VZGC008	-3.68	VCGC008	-27.39
VZGC009	-7.05	VCGC009	-19.55
VZGC010	-5.96	VCGC010	-22.44
VZGC011	-10.89	VCGC011	-19.46
VZGC012	-4.29	VCGC012	-23.28
VZGC013	-4.25	VCGC013	-24.72
VZGC014	-3.89	VCGC014	-20.55
VZGC015	-13.79	VCGC015	-32.94
VZGC016	-9.25	VCGC016	-13.21
VZGC017	-1.14	VCGC017	-19.28
VZGC018	-0.26	VCGC018	2.23
VZGC019	-4.93	VCGC019	-7.24
VZGC020	-8.69	VCGC020	-15.50
VZGC021	-6.59	VCGC021	-35.98
VZGC022	-4.17	VCGC022	-29.93
VZGC023	-7.75	VCGC023	-26.77
VZGC024	-6.91	VCGC024	-40.83
VZGC025	-3.15	VCGC025	-7.93
VZGC026	-4.17	VCGC026	-29.65
VZGC027	-0.63	VCGC027	7.59
VZGC028	0.16	VCGC028	-28.22
VZGC029	-1.65	VCGC028	-41.57
VZGC030	-5.83	VCGC030	-34.31

Table V.1. Some high scoring peptide fragments selected from BMP-2 region that binds to noggin. Their Glide (GlS) and GOLD (GoS) scores along with noggin residues with which they interact in binding are given.

Pep	BMP-2	Noggin	GlS	GoS	Residues of noggin in close contact with peptides after docking
Pep1	V33,A34,P35, P36	46-47,49,206,208-210	-2.2	37.9	H49, F54, R204, R206, C207
Pep1.1	A34, P35	46-47,49	-2.1	-	D45, E48, D163, G165
Pep1.2	A34, P35, P36	46-47,49210	-5.2	38.4	L46, F54, R204, C207
Pep1.3	P35, P36	46-47,49,210	-3.8	31.2	L46, F54
Pep1.4	V33, A34	46-47,49,206,208-210	-4.9	29.3	L46, F54, R204
Pep1.5	V33, A34, P35	46-47,49,206,208-210	-4.9	33.1	L46, F54, R204, C207
Pep1.6	V33, P36	47,49,206,208-210	-3.5	28.4	L46, H49, F54
Pep2	P50,L51, A52, D53	27,30-33	-4.1	27.6	M27, Q28, Y30, L31, H32
Pep2.1	P50, L51	30-33	-4.7	14.7	H29, L31, I33, R34, A36
Pep2.2	L51, A52	30-33	-4.2	-	H29, R34, A36, P37
Pep2.3	A52, D53	27,30-32	-4.4	41.1	H29, L31, R34, A36
Pep2.4	P50, A52	30-33	-4.6	14.1	H29, L31, I33, R34, A36
Pep2.5	P50, D53	27,30-33	-4.6	24.0	H29, L31, I33, R34, A36
Pep2.6	L51, D53	27,30-33	-3.0	23.6	H29, I33, R34, A36
Pep3	E96, K97, V98	204,206,208,218-221	-4.7	51.5	D39, L43, L46, H199, I220, Q221, P223
Pep3.1	E96, K97	207,208,218-221	-2.8	18.1	L43, L46, I220, Q221, Y222, P223
Pep3.2	K97, V98	204,218-221	-5.0	-	P37, D39, L43, L46, Q221, P223
Pep3.3	E96, V98	204,206,208218-221	-3.3	34.5	F54, R204, R206, C207
Pep4	V99, L100, K101	37-40,43,199,220-221,223	-3.9	41.2	D39, L43, L46, H199, I220, Q221, P223
Pep4.1	V99, L100	199,220-221,223	-3.9	-	D39, N40, K193, P194, S195, S197
Pep4.2	V99, K101	37-40,43,199,220-221,223	-0.9	39.8	P37, S38, D39, L43, L46, P223
Pep4.3	L100, K101	37-40,43,220-221,223	-5.4	40.3	P37, D39, L43, L46, H199, Q221, P223
Pep5	N102, Y103, Q104	35-40,43	-0.5	38.7	P37, S39, D39, L43, L46, H199, I220, Q221, P223
Pep5.1	N102, Y103	35-40,43	-6.5	39.5	D39, L43, L46, I220, Q221, P223
Pep5.2	Y103, Q104	35-40	-4.2	23.5	P37, S38, D39, L43, S197, H199, P223
Pep5.3	N102, Q104	35-40,43	-3.4	36.3	P37, S38, D39, H199, P223

7.0 FIGURES

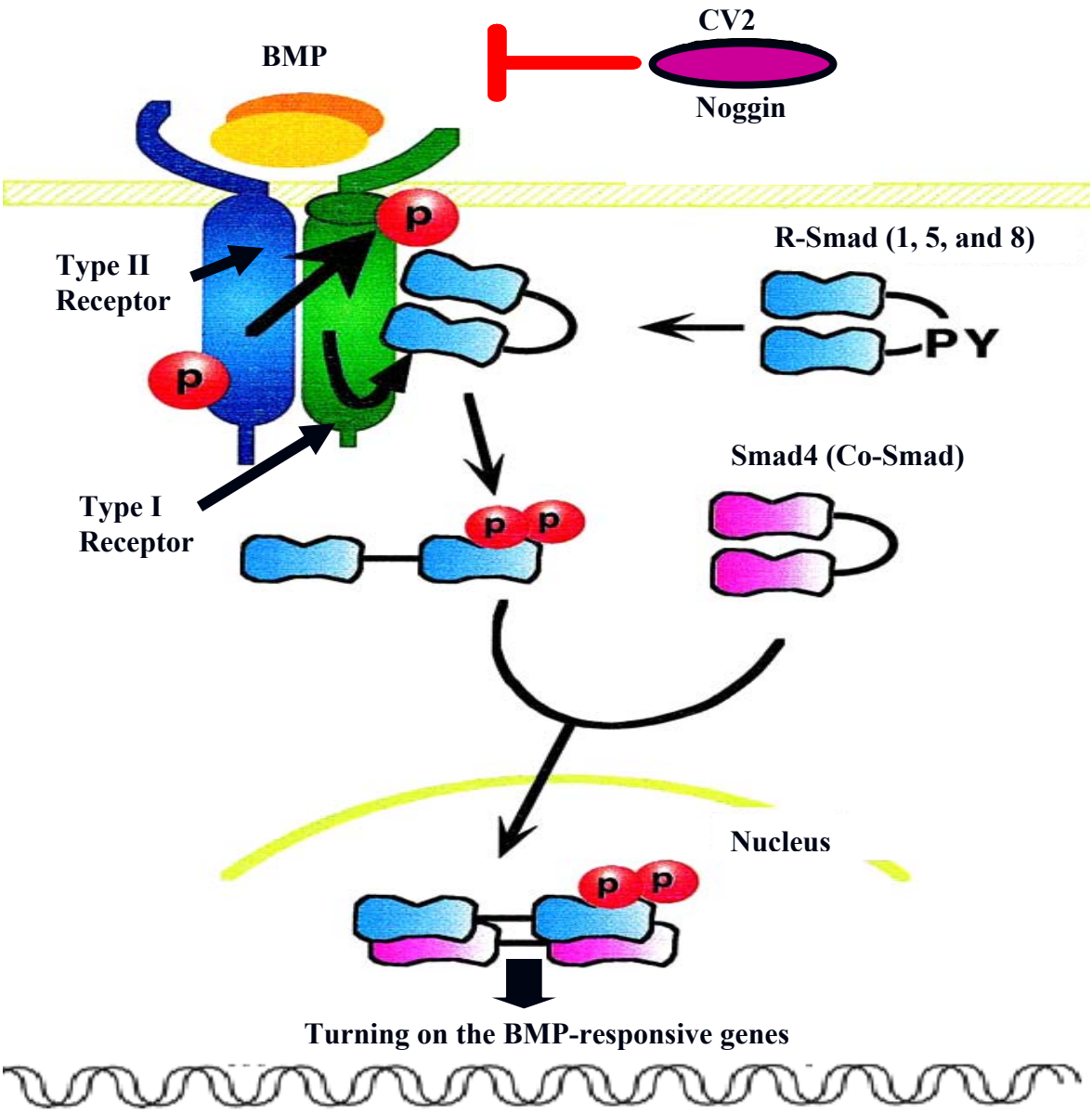

Figure 1.1. BMP signaling pathway. BMP dimer binds to BMP receptor type II. This binding recruits type I receptors, so that a hetero-tetramer is formed with two receptors of each type. The proximity of the receptors allows the type II receptor to phosphorylate the type I receptor. Activated type I receptors phosphorylate R-Smads or receptor regulated Smads (Smad1/5/8) which form complexes with Smad4. Activated Smad complexes regulate gene expression of several target genes. The inhibitory Smads (Smad6, Smad7) antagonize signaling. BMP can also be inhibited by antagonist noggin or CV2.

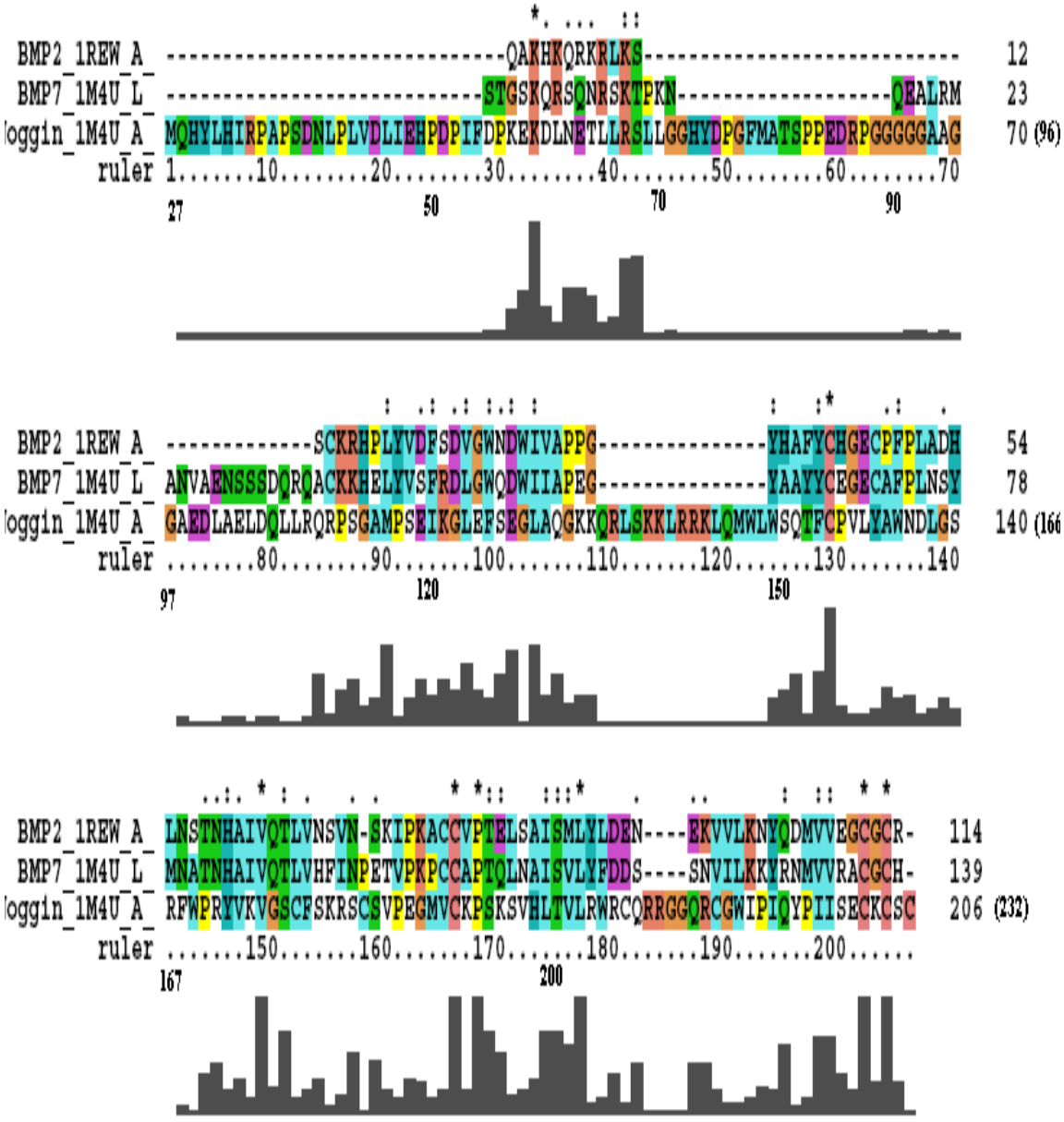

Figure 1.2. Sequence alignment performed using ClustalW: Multiple sequence alignment of BMP-2 (1rew), BMP-7 (1m4u), and noggin (1m4u) sequences to identify equivalent residues among these sequences.

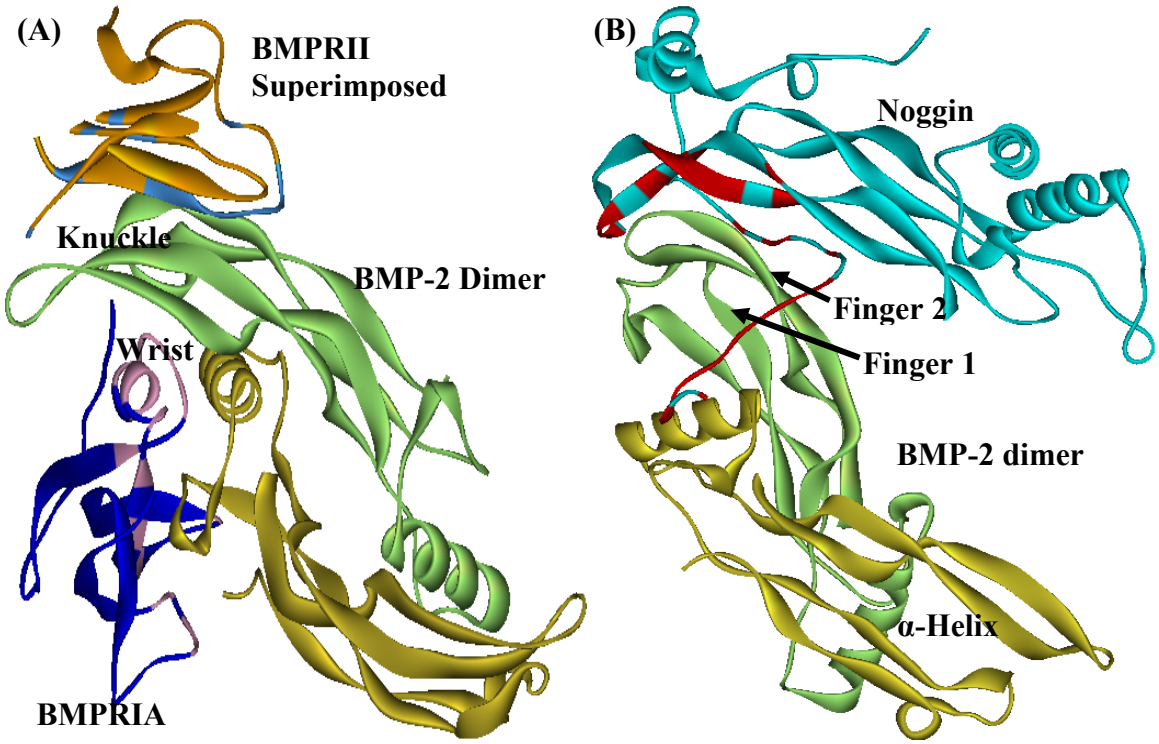

Figure 1.3. Model structures of complexes: (A) Model ternary complex of BMP-2/BMPRIA/BMPRII showing BMP-2 dimer binding to BMPRIA ectodomain (PDB ID 1rew) and BMPRII. Color pink and light blue show the BMP-2 binding regions of BMPRIA ectodomain and BMPRII, respectively. (B) Model structure of BMP-2 dimer binding to the inhibitor noggin based on BMP-7/noggin dimer (1m4u). The binding regions of noggin to the BMP-2 are colored in red.

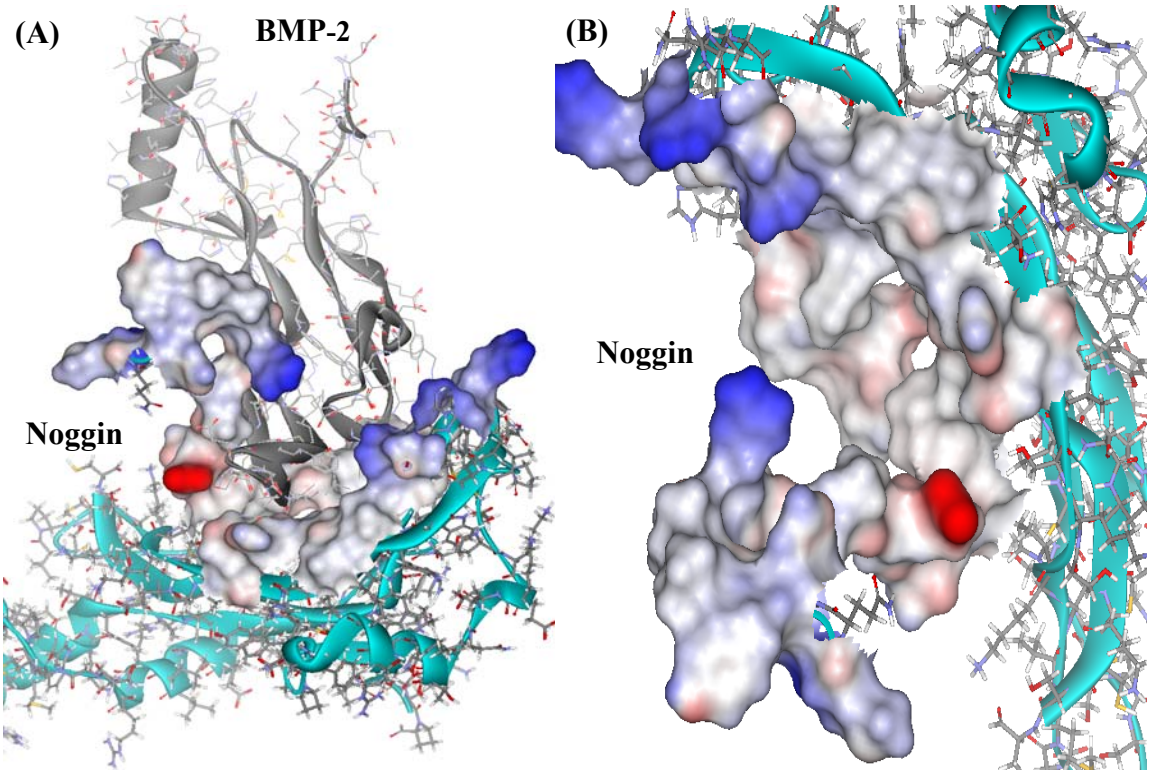

Figure 1.4. Noggin surface involved in binding with BMP-2. (A) Both noggin and BMP-2 are shown with solid surface of the noggin residues involved in binding. (B) A close snapshot of the noggin binding surface showing several interesting concave groves with hydrophobic and hydrophilic hotspots where small molecules can find tight binding spots.

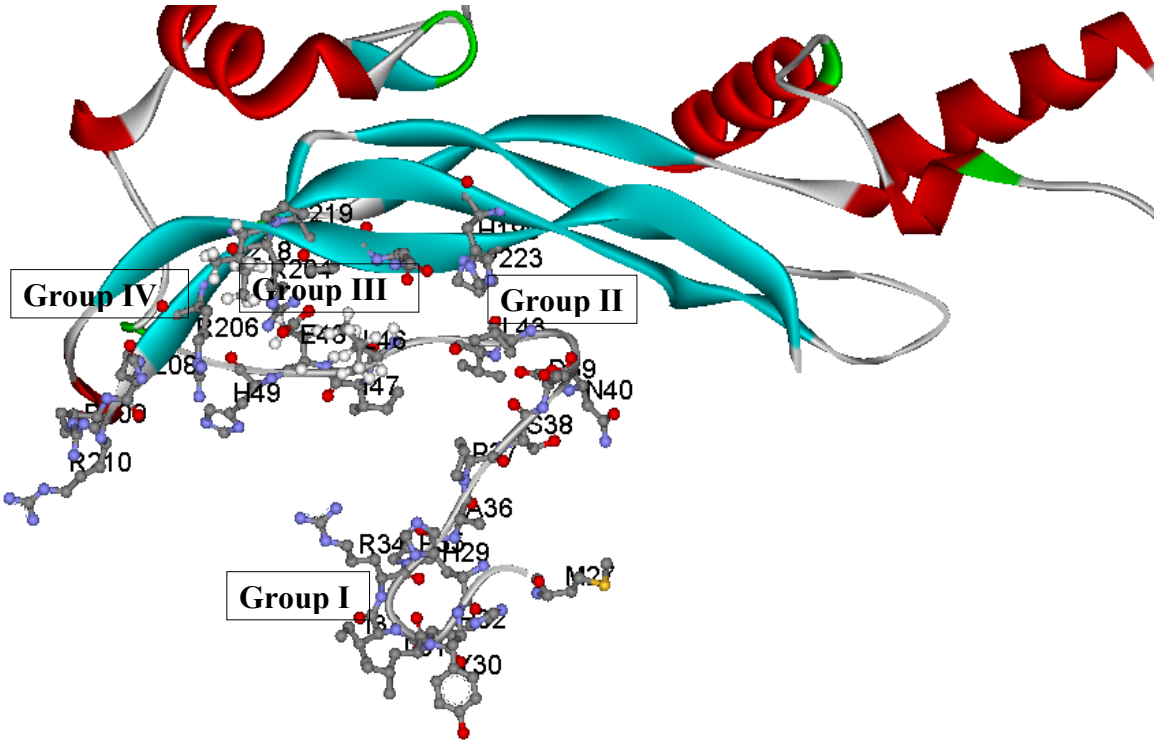

Figure 1.5. Noggin binding regions to BMP are divided into four groups for the LUDI docking. Group I consisting of AAs at positions 27, 29-37, Group II consisting of AAs at positions 37-46, 223, Group III consisting of AAs at positions 47, 48, 199, 204, 218-221 and Group IV consisting of AAs at positions 206-210.

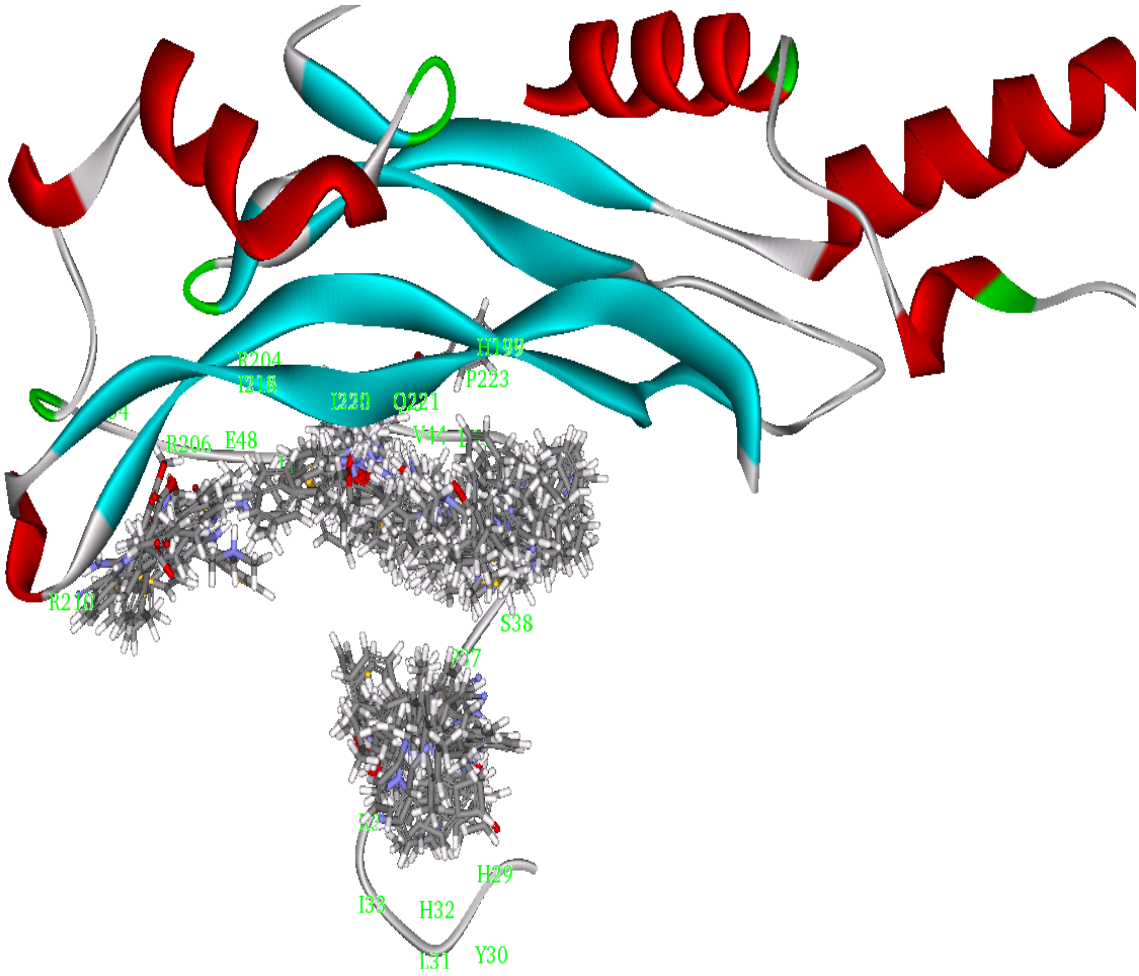

Figure 1.6. Distribution of the top scoring 100 molecules by the LUDI docking over the entire binding region of noggin showing at least three hotspots where most of the top scoring molecules find suitable binding pockets.

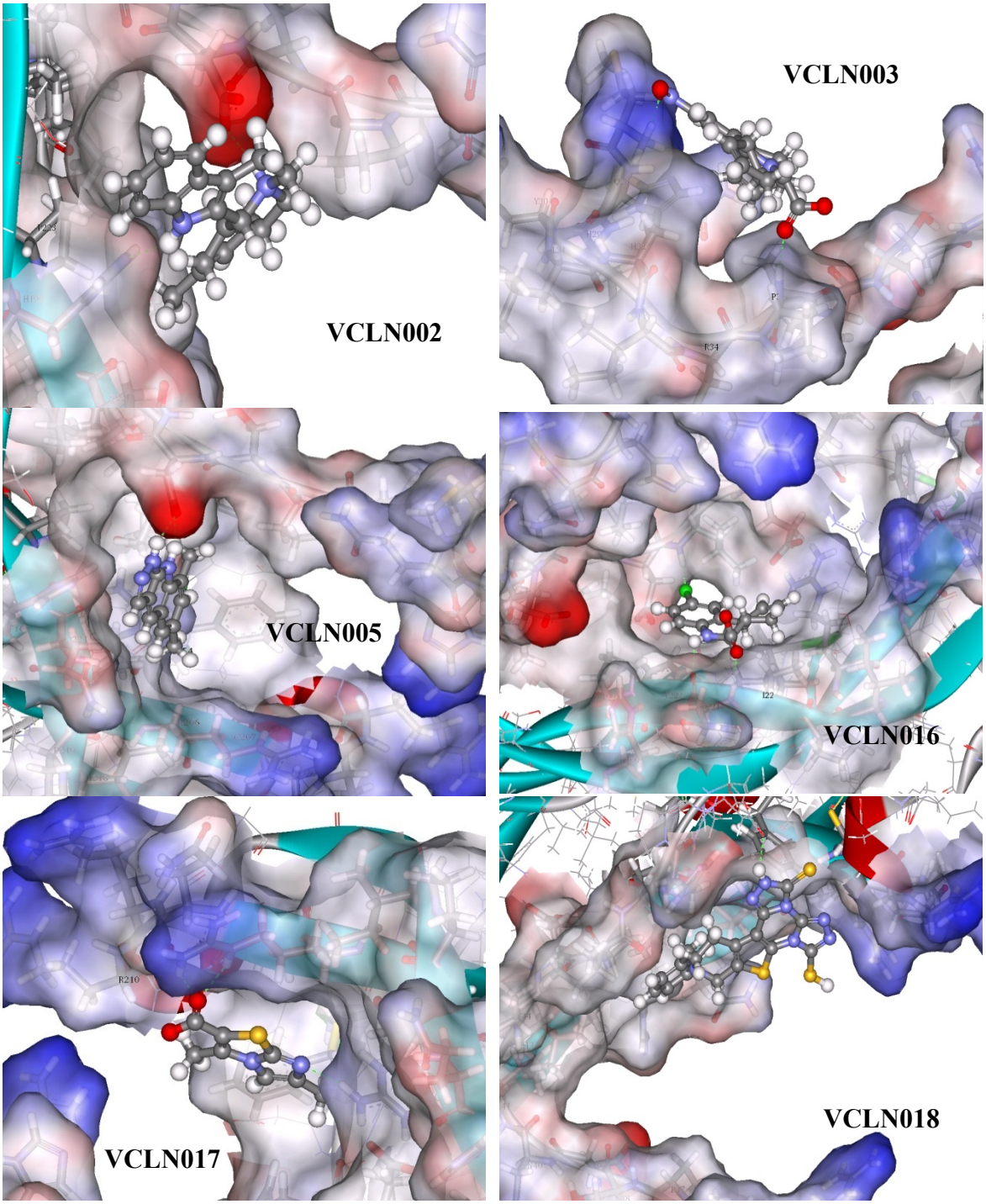

Figure 2.1. Docking poses of some selected top scoring molecules from Ludi3 scores using the CAP database.

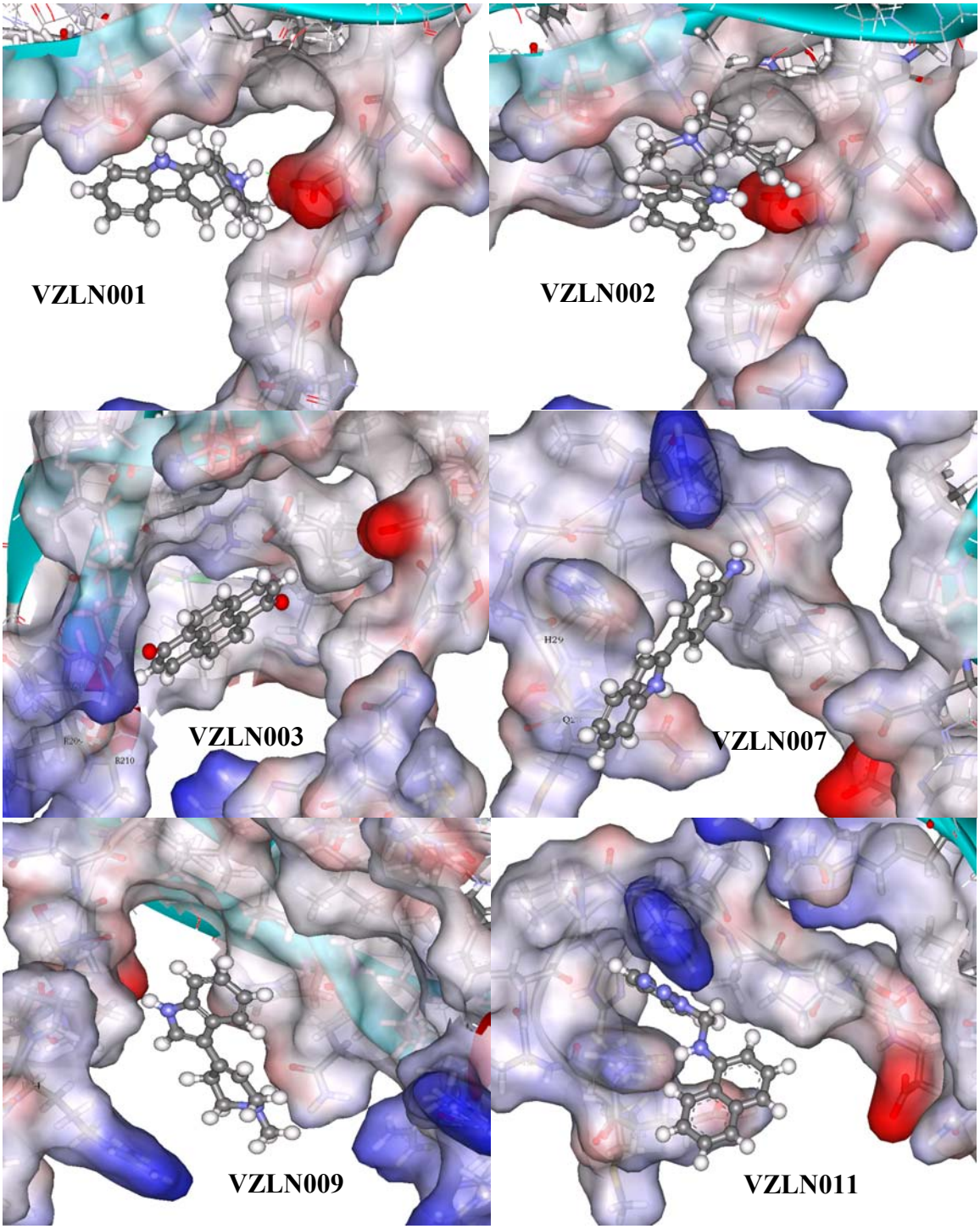

Figure 2.2. Docking poses of some selected top scoring molecules from Ludi3 scores using the ZINC database.

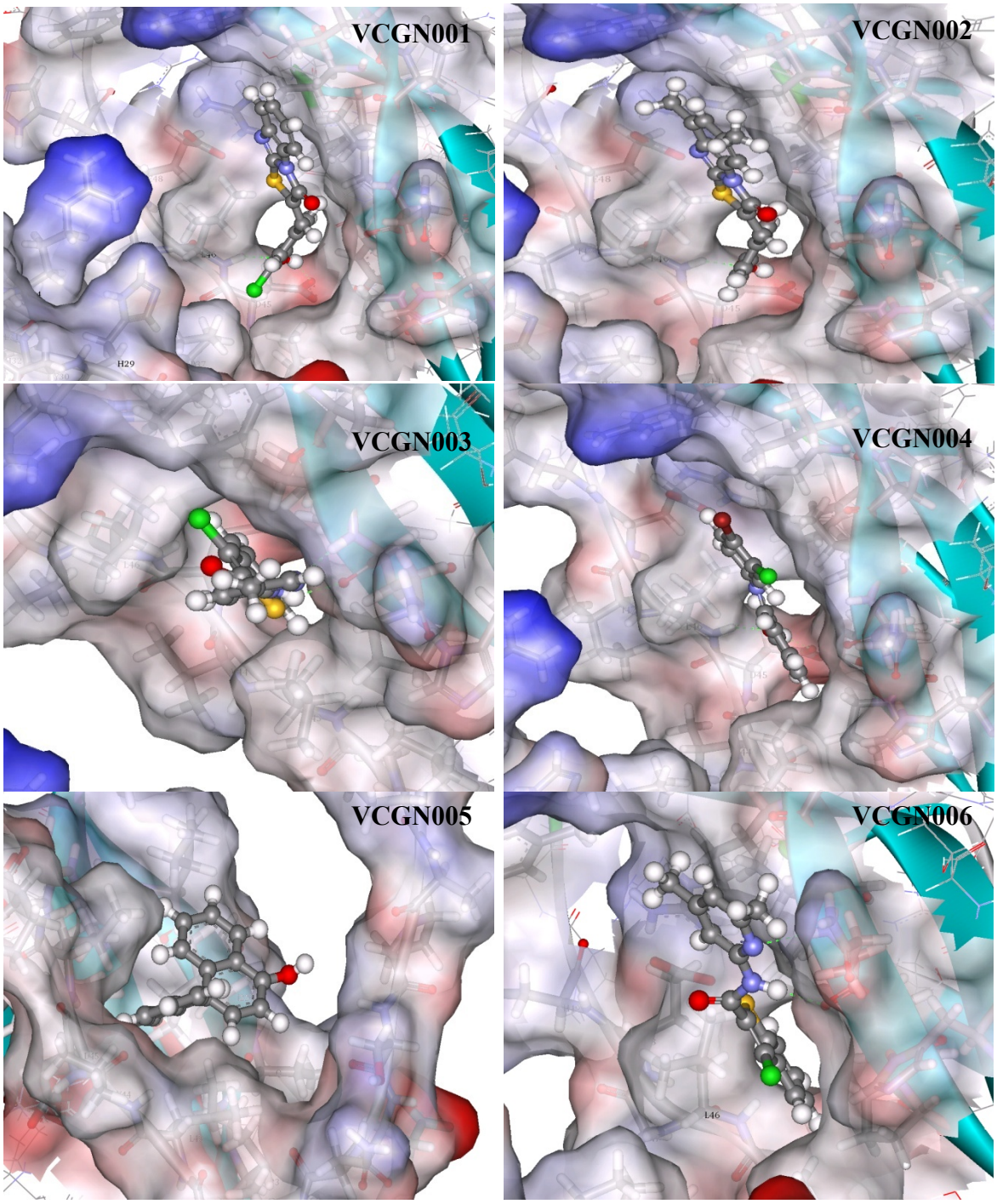

Figure 2.3. Docking poses of some selected top scoring molecules from Glide, GOLD, and CombScore using the CAP database.

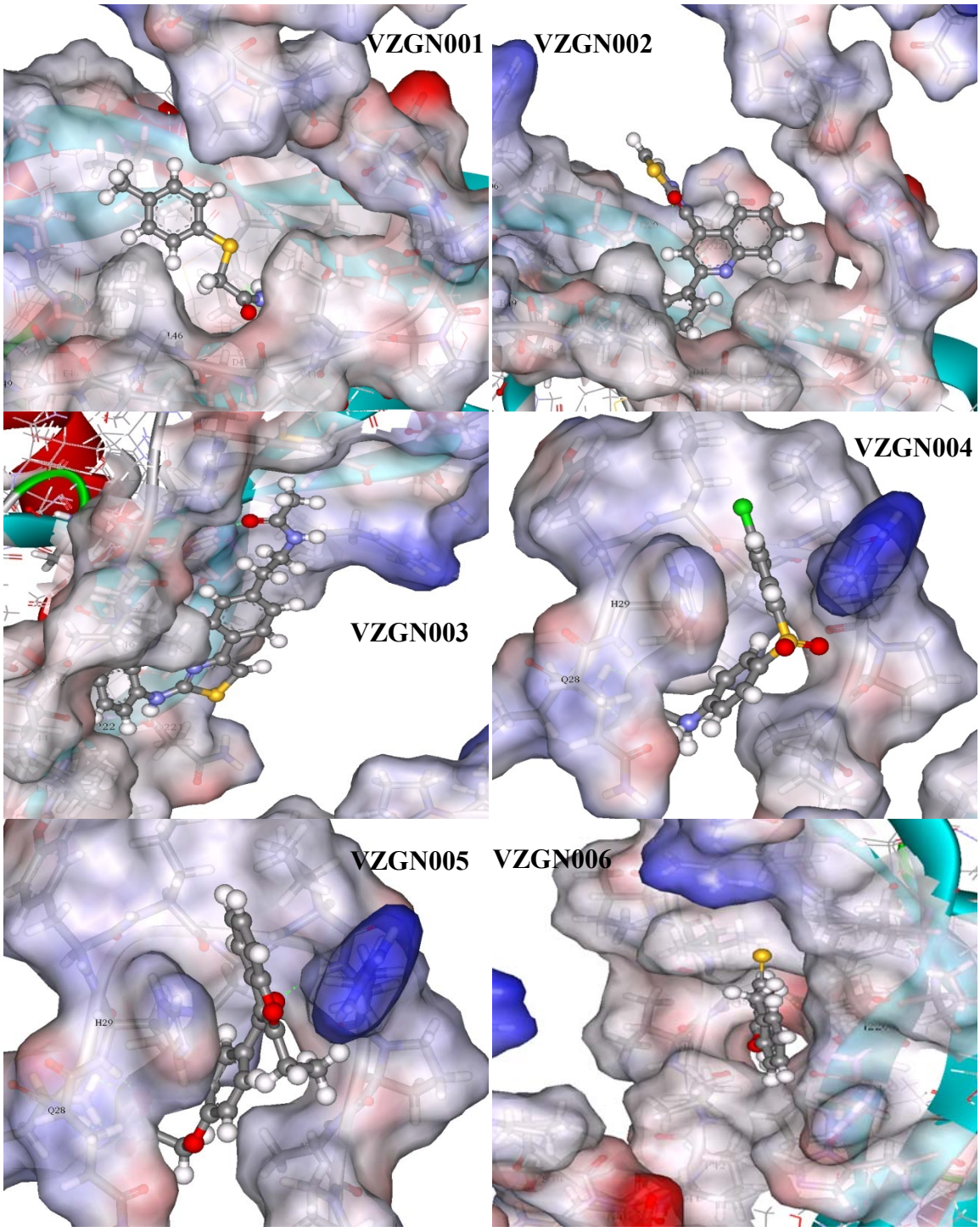

Figure 2.4. Docking poses of some selected top scoring molecules from Glide, GOLD, and CombScores using the ZINC database.

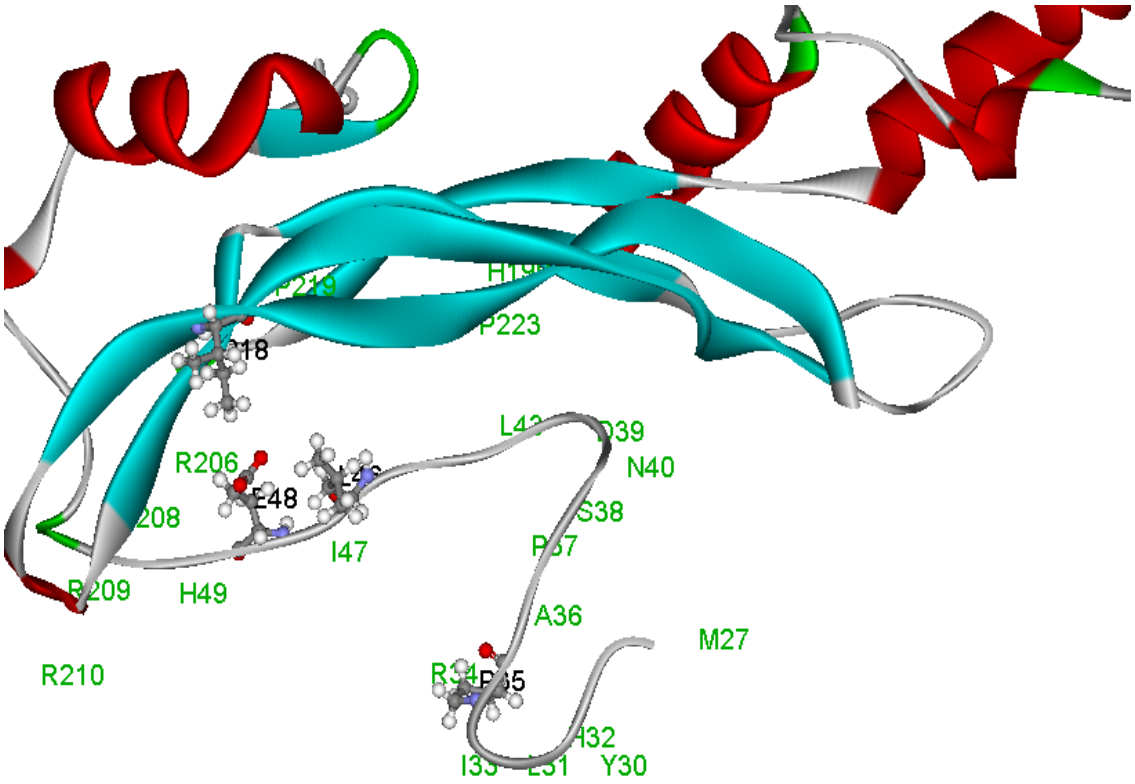

Figure 2.5. Solid ribbon backbone trace of noggin showing its binding residues to the BMP dimer. The AAs that are in contact with BMP are labeled. Mutational studies of four AAs, L46D, E48K, I218E and P35R, of noggin, shown in ball and stick, were reported to cause loss of binding affinity to BMP, thus suggesting these could be important target sites.

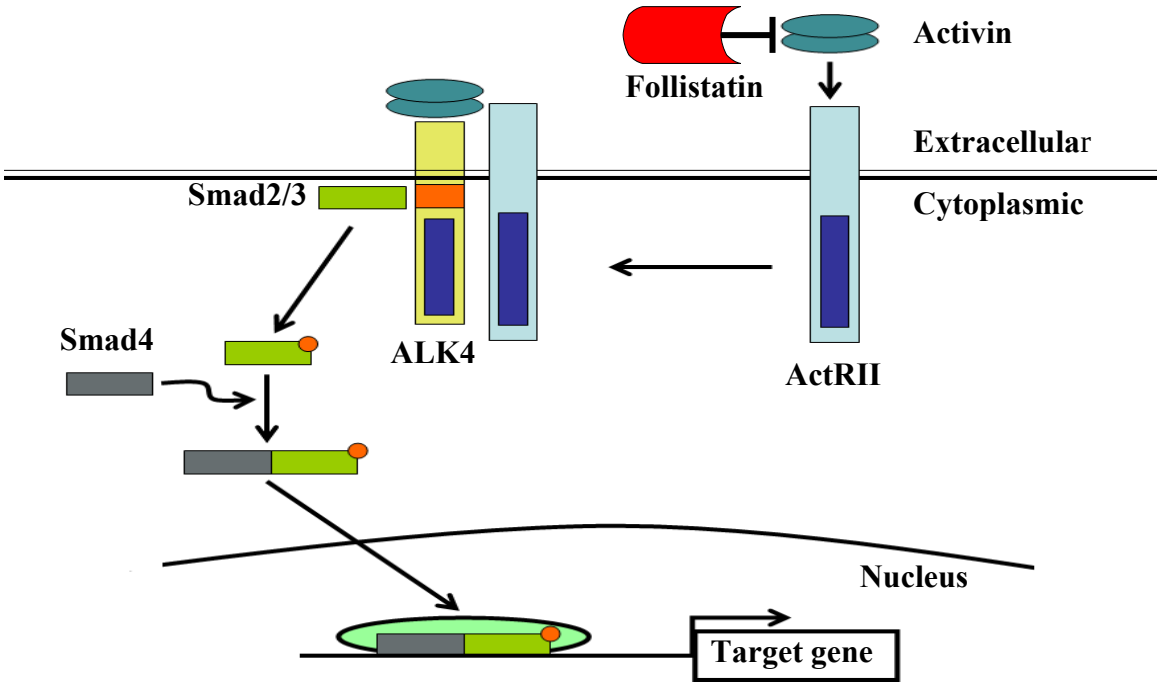

Figure 3.1. Activin signaling pathway: Activin dimer binds to activin type II receptor (ActRIIB). This binding recruits type I receptors (ALK4) so that a hetero-tetramer is formed with the two receptors of each type. The proximity of the ActRII helps to phosphorylate ALK4 which in turn phosphorylates receptor regulated R-Smads (Smad 2/3). The R-Smad complexes with common Smad (Smad4) and these activated Smad complexes regulate gene expression of several target genes after translocating into nucleus. The activin at extra cellular level is inhibited (regulated) by antagonist follistatin.

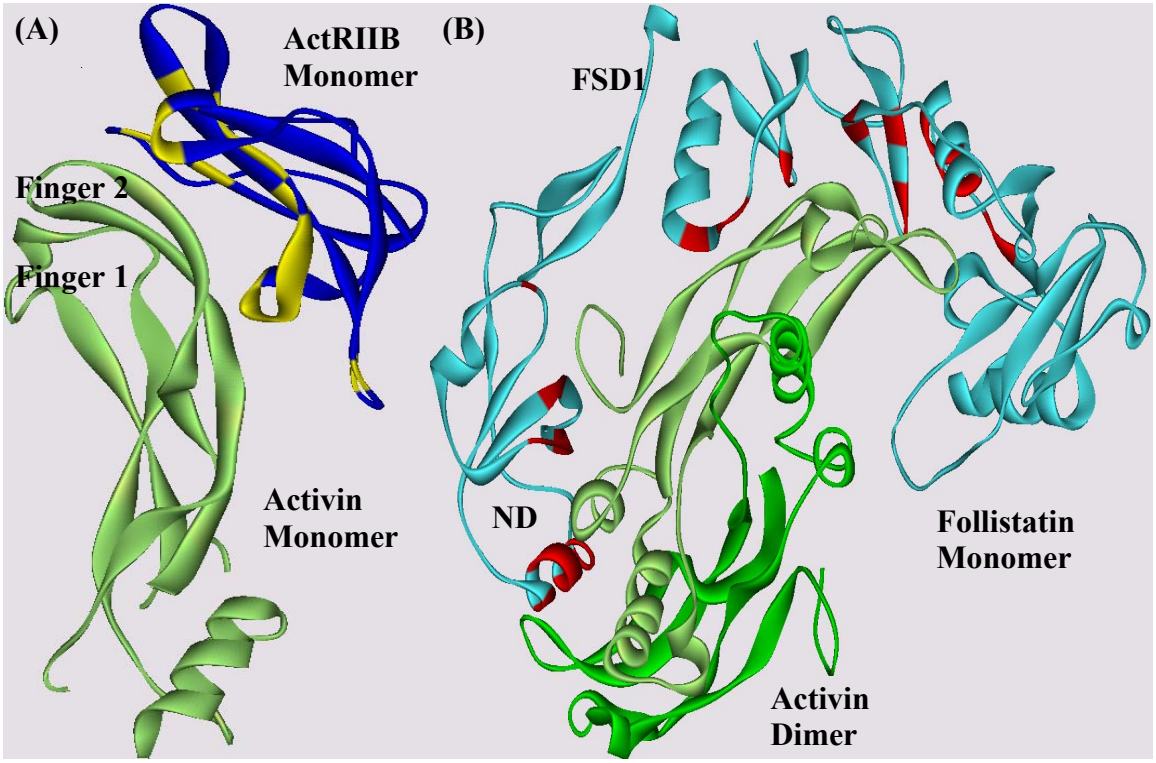

Figure 3.2. Crystal structures of activin complexes: (A) Complex showing activin binding to ActRIIB extracellular domain (PDB ID: 1s4y). The binding regions of the ActRIIB extracellular domain to the activin are colored in yellow. (B) Structure of activin dimer binding to the antagonist follistatin (2b0u). The binding regions of follistatin to the activin are colored in red.

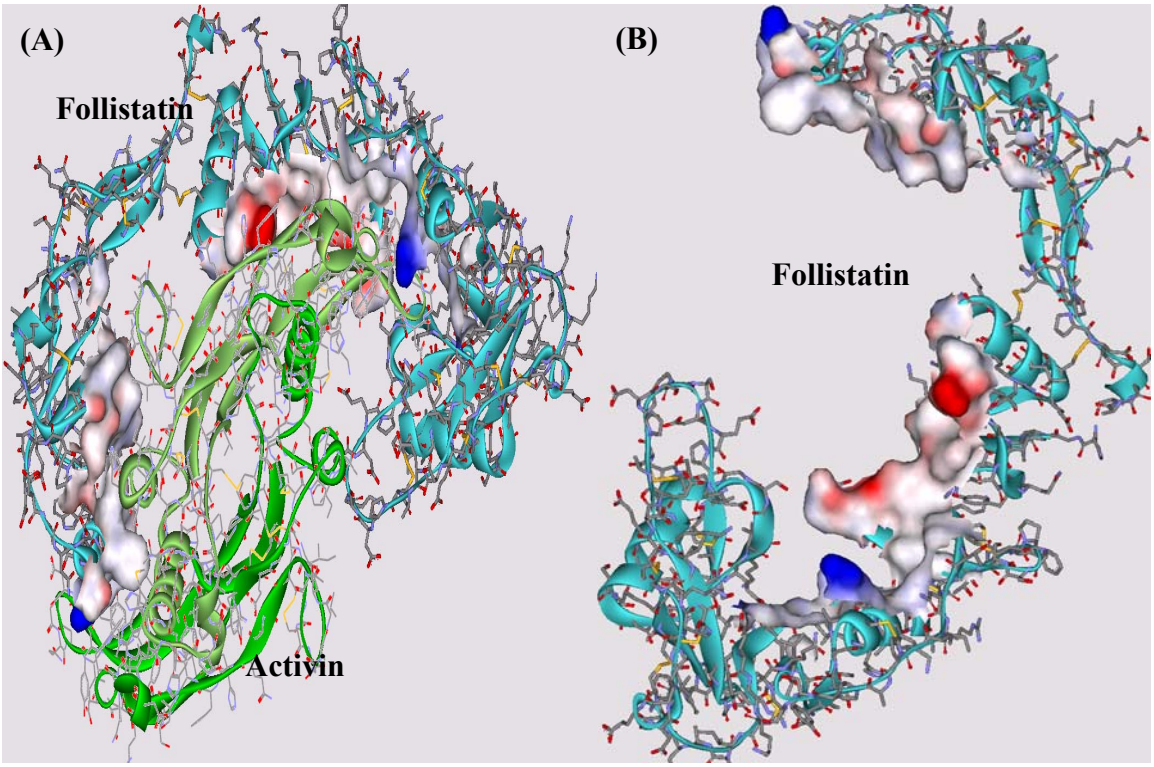

Figure 3.3. Follistatin/activin complex: (A) Follistatin shown with solid surface of its residues involved in binding with activin. (B-F) Close snapshots of the follistatin binding surface showing several interesting concave groves with hydrophobic and hydrophilic hotspots where small molecules can find tight binding spots.

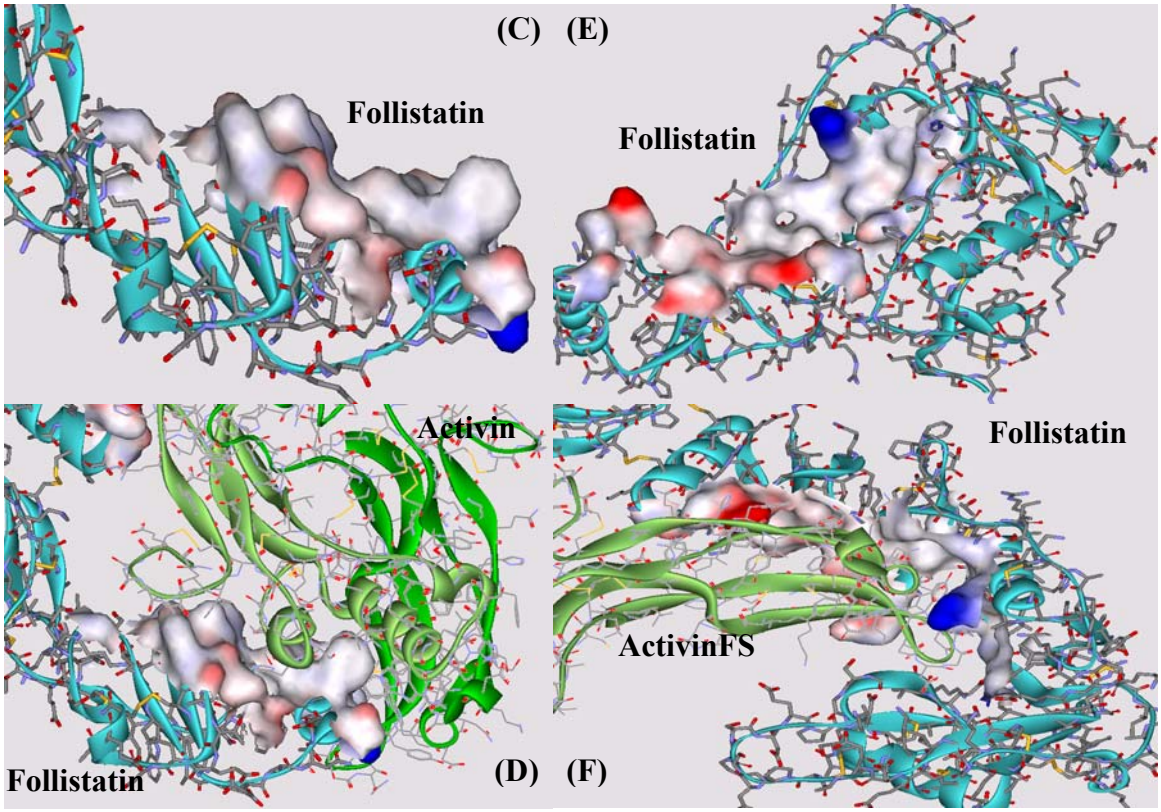

Figure 3.3. (Continued…)

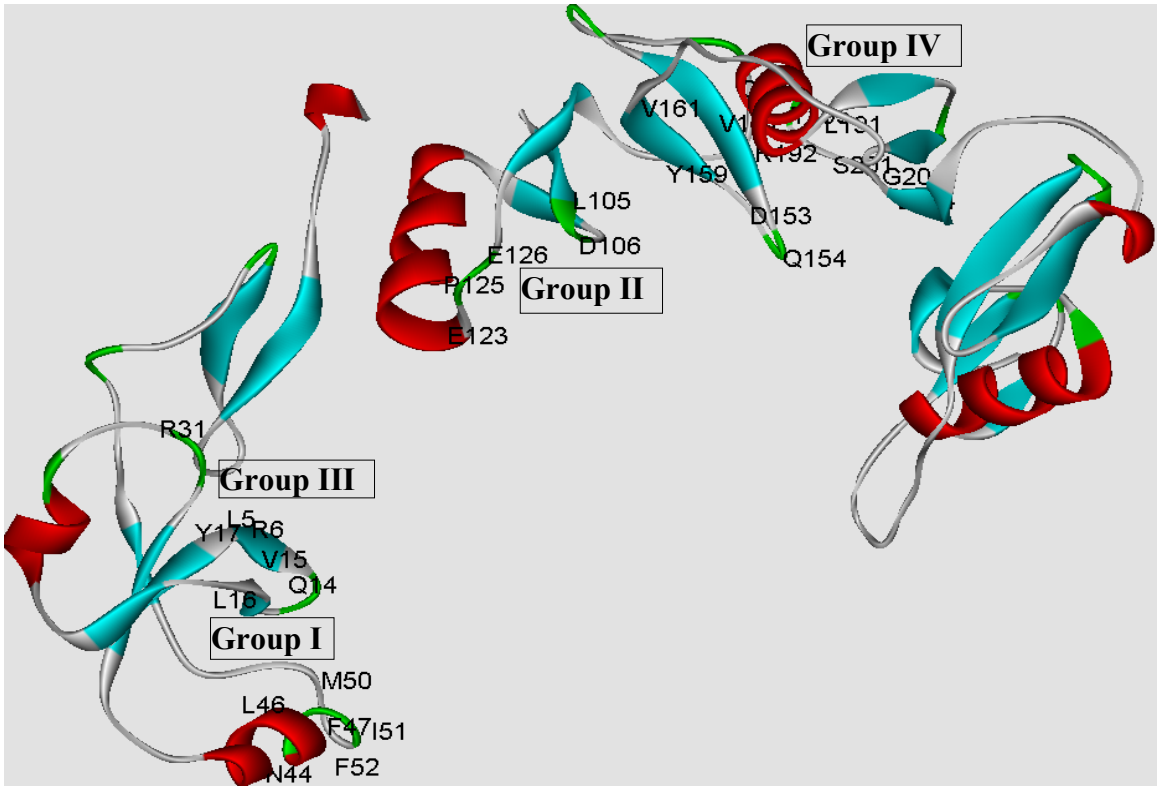

Figure 3.4. Follistatin binding regions to activin are divided into four groups. Group I consisting of amino acids (AAs) at positions 5, 6, 14-17, 44, 46-48, 50-52; Group II consisting of AAs at positions 105, 106, 123, 125, 126, and 159; Group III consisting of AAs at positions 5, 6, and 31; and Group IV consisting of AAs at positions 151, 153, 154, 161, 191, 192, 195, 196, 201, 203, and 204.

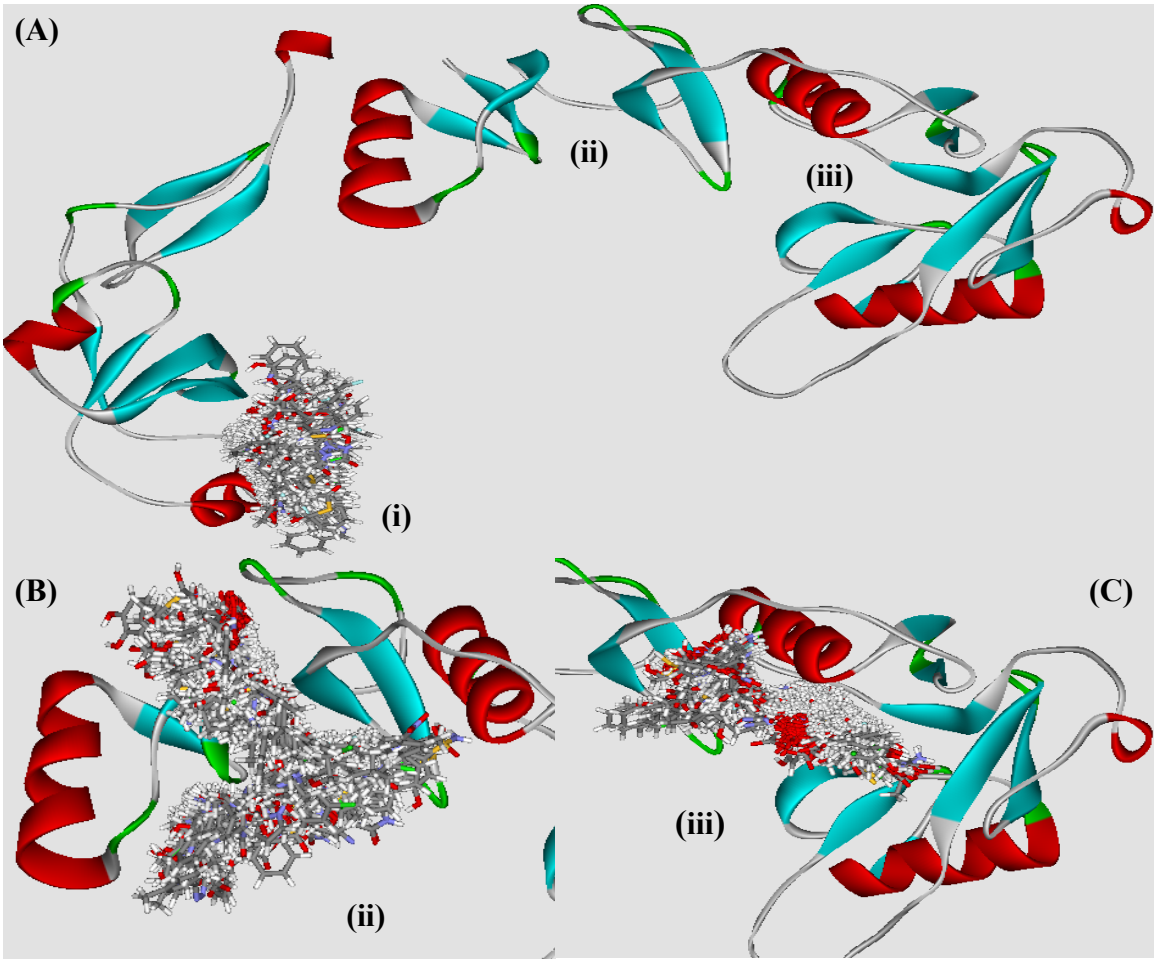

Figure 3.5. Distribution of top scoring 150 molecules by the LUDI docking over the binding region of follistatin. (A) Follistatin showing at least three regions (i), (ii) and (iii) where most of the top scoring molecules find suitable binding hotspots. A set of top scoring 50 molecules clustered at binding hotspot in region (i). (B) Showing top scoring 50 molecules clustered at binding hotspot in region (ii), and (C) showing top scoring 50 molecules clustered at another binding hotspot in region (iii).

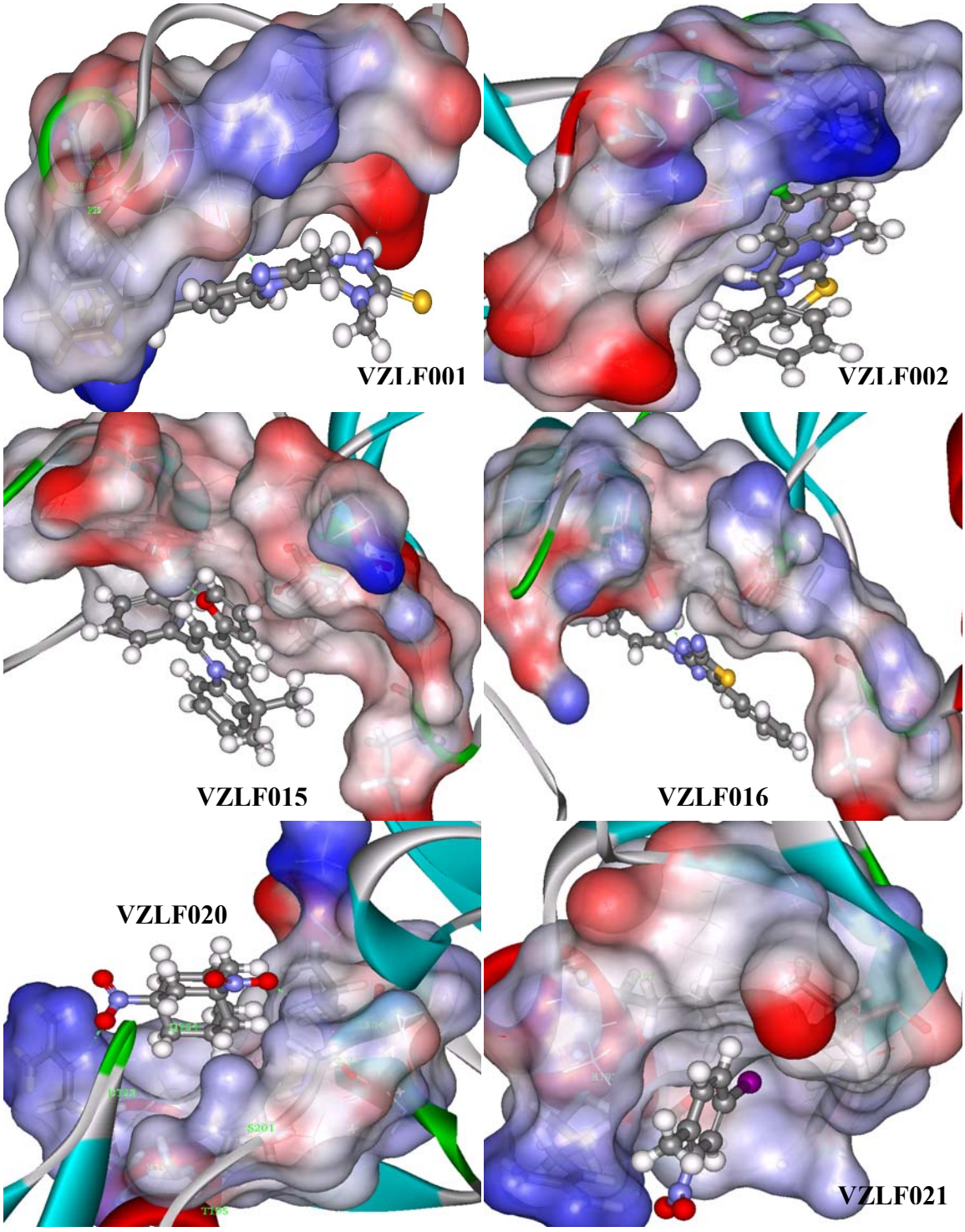

Figure 3.6. Docking poses of few top scoring small molecules using the LUDI *de novo* design docking procedure using the ZINC database - two molecules from each hotspot region are shown.

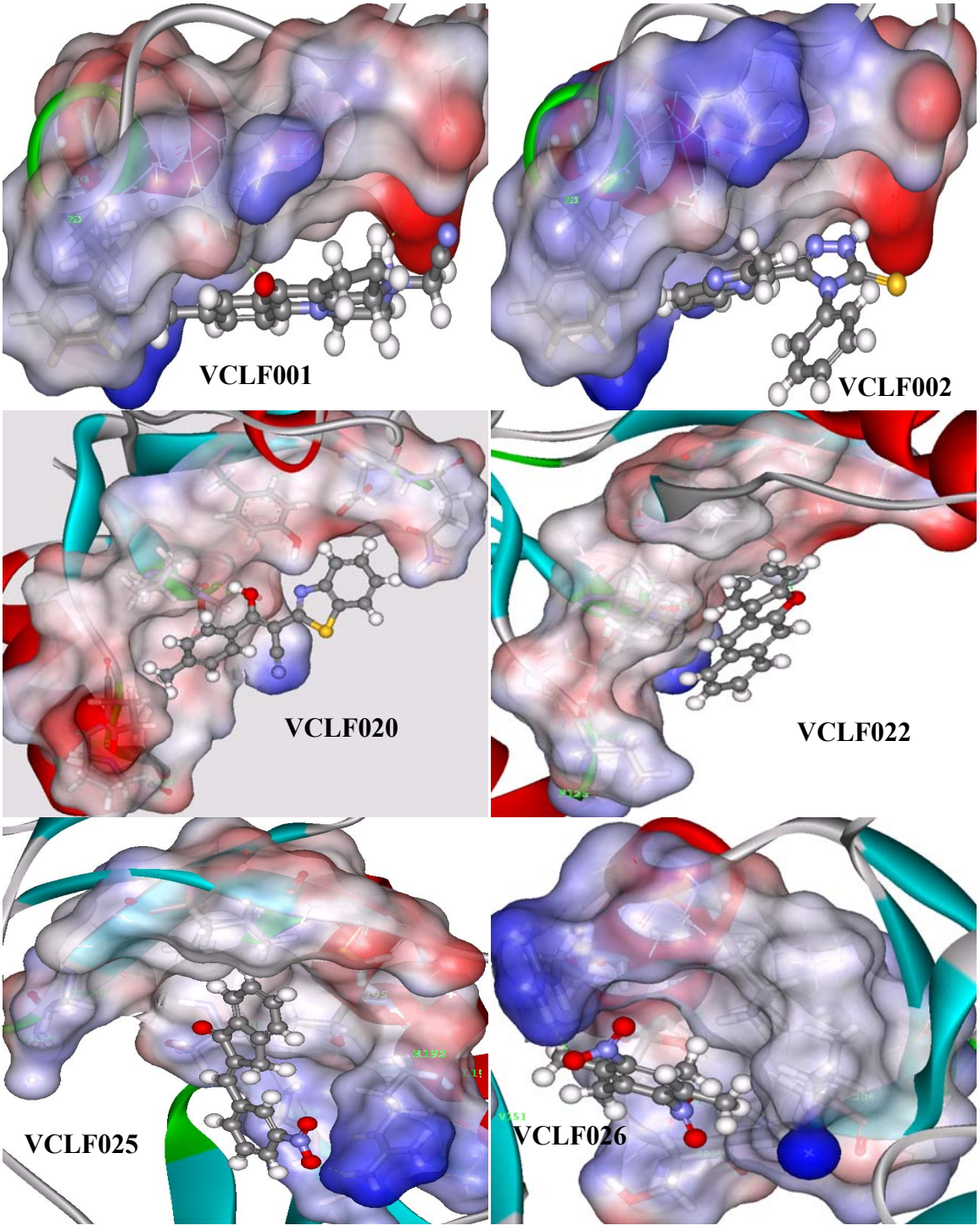

Figure 3.7. Docking poses of few top scoring molecules using LUDI de novo design docking procedure using the CAP database: two from each hotspot region.

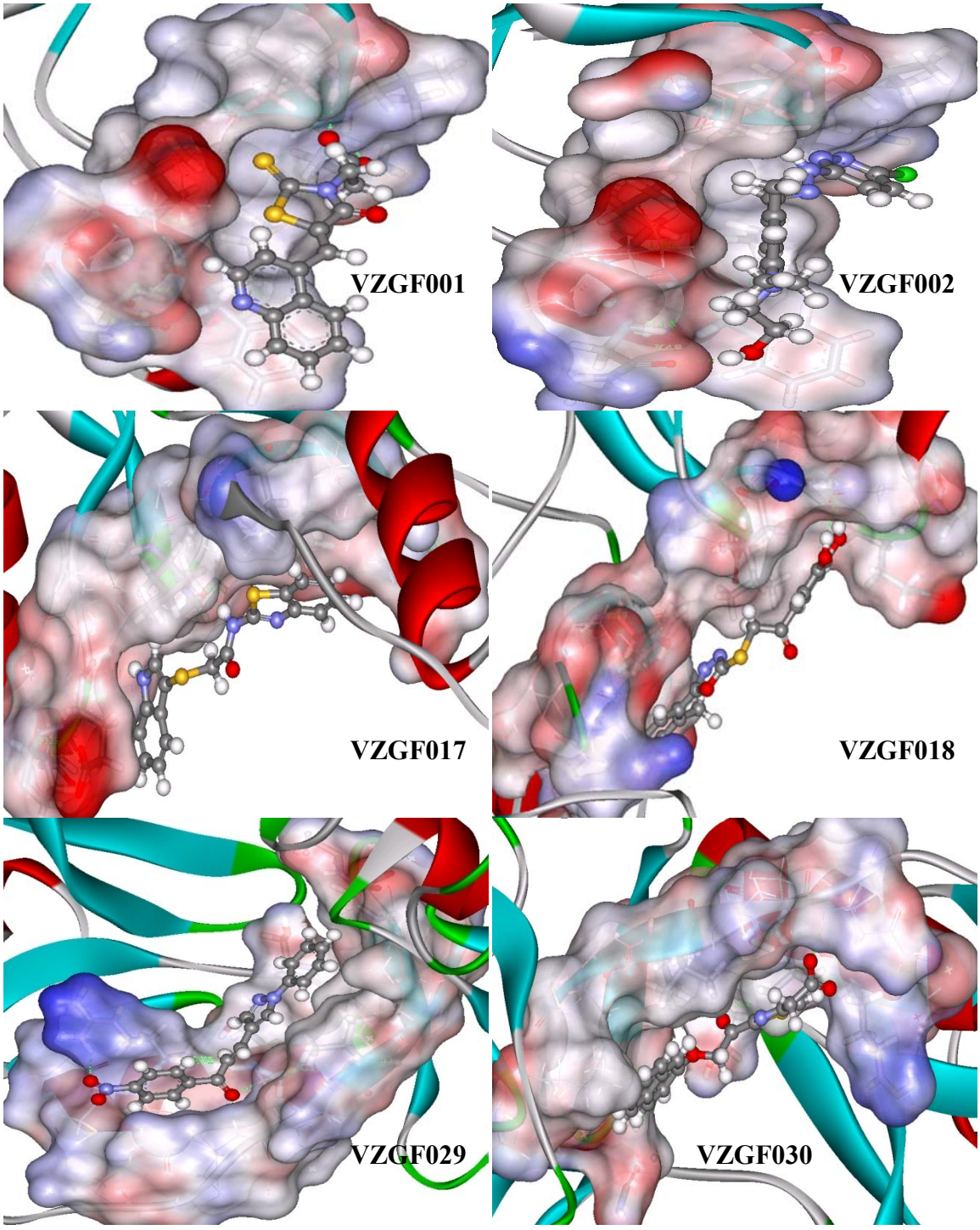

Figure 3.8. Docking poses of few selected top scoring small molecules from ZINC database using the Glide-XP and the GOLD docking scores - two molecules from each hotspot region are shown.

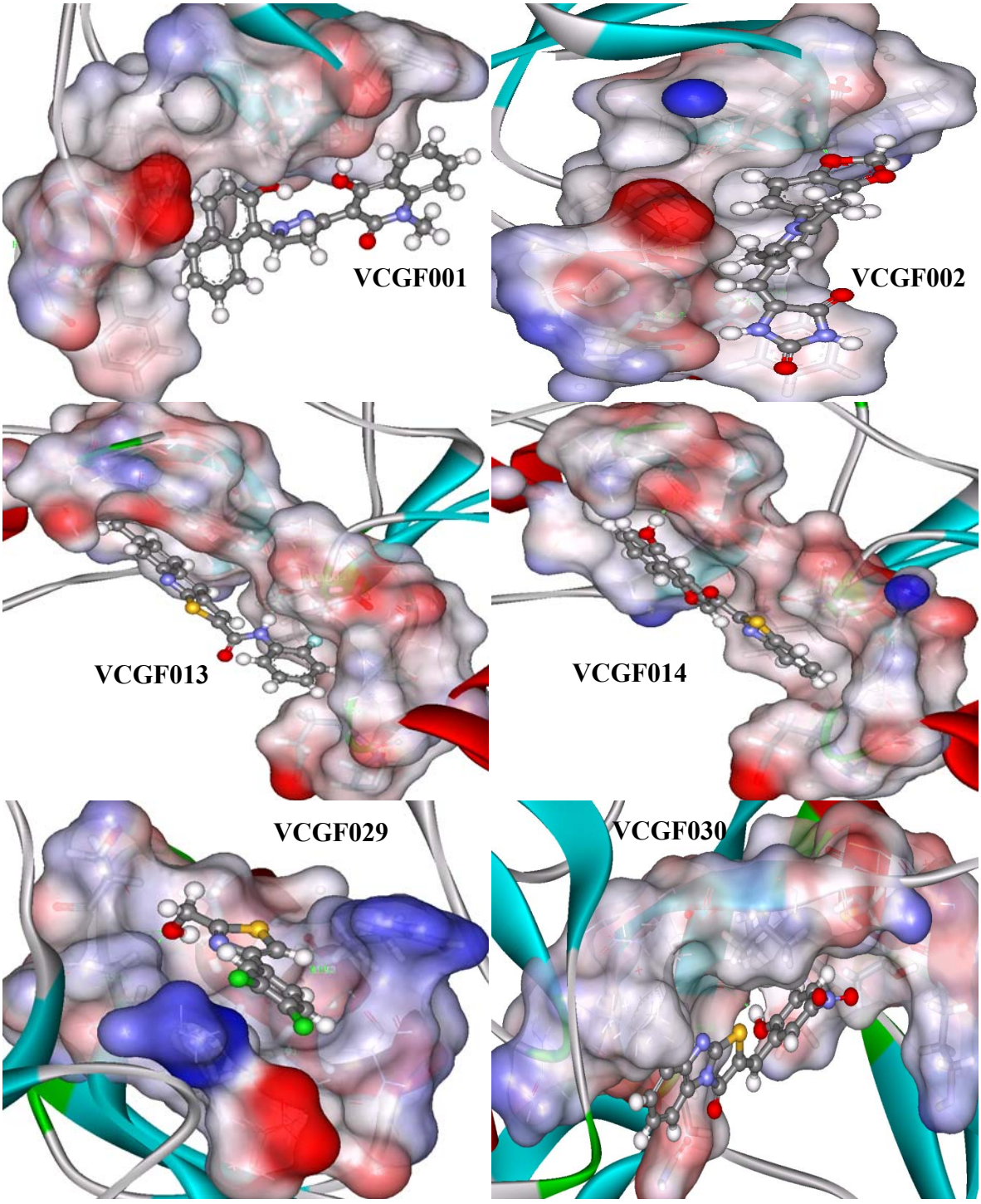

Figure 3.9. Docking poses of few top scoring small molecules from the CAP database using the Glide-XP and GOLD docking procedures - two molecules from each hotspot region are shown.

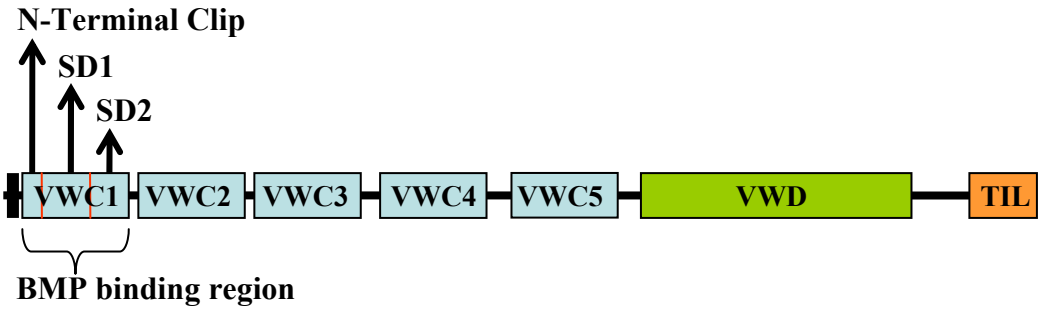

Figure 4.1. A schematic representation of CV2 domain architecture along its sequence of 668 amino acids.

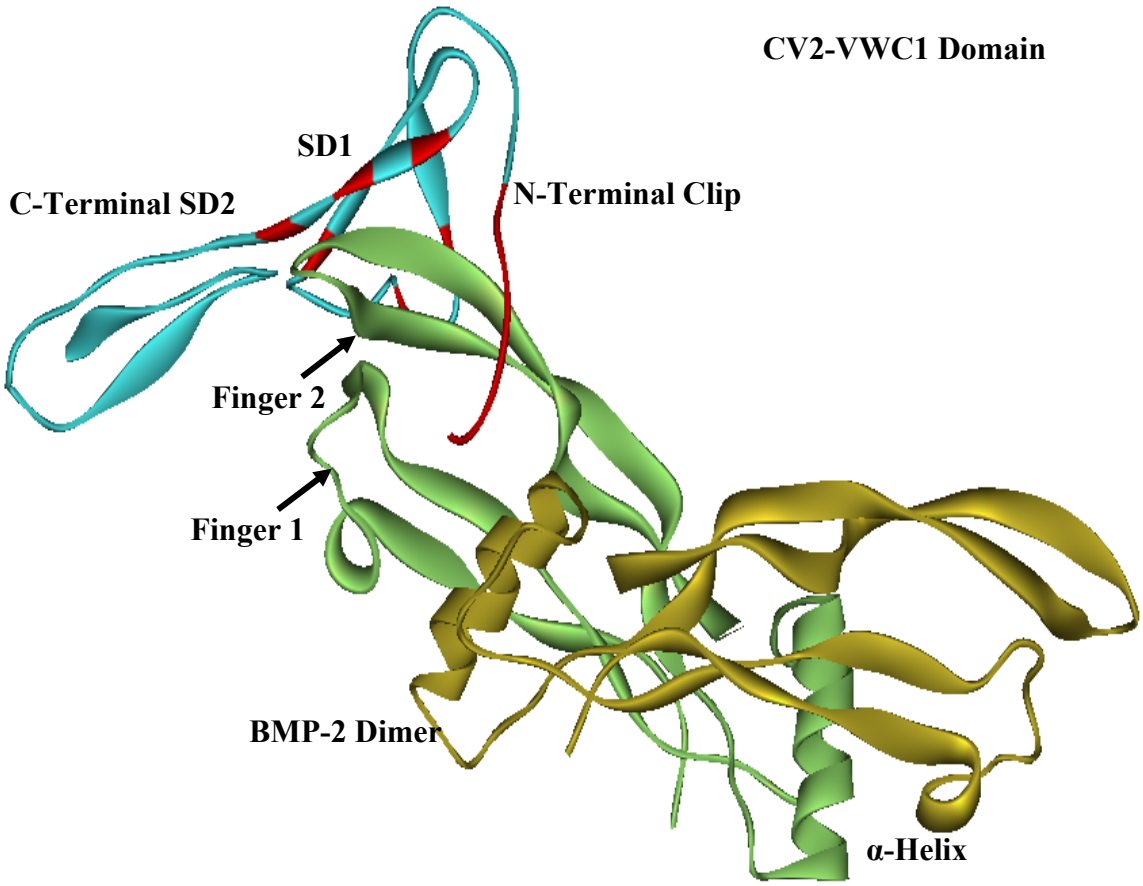

Figure 4.2. Structure of BMP-2 dimer binding to the inhibitor CV2-VWC1 (3bk3). BMP binding regions of CV2-VWC1 are shown in red.

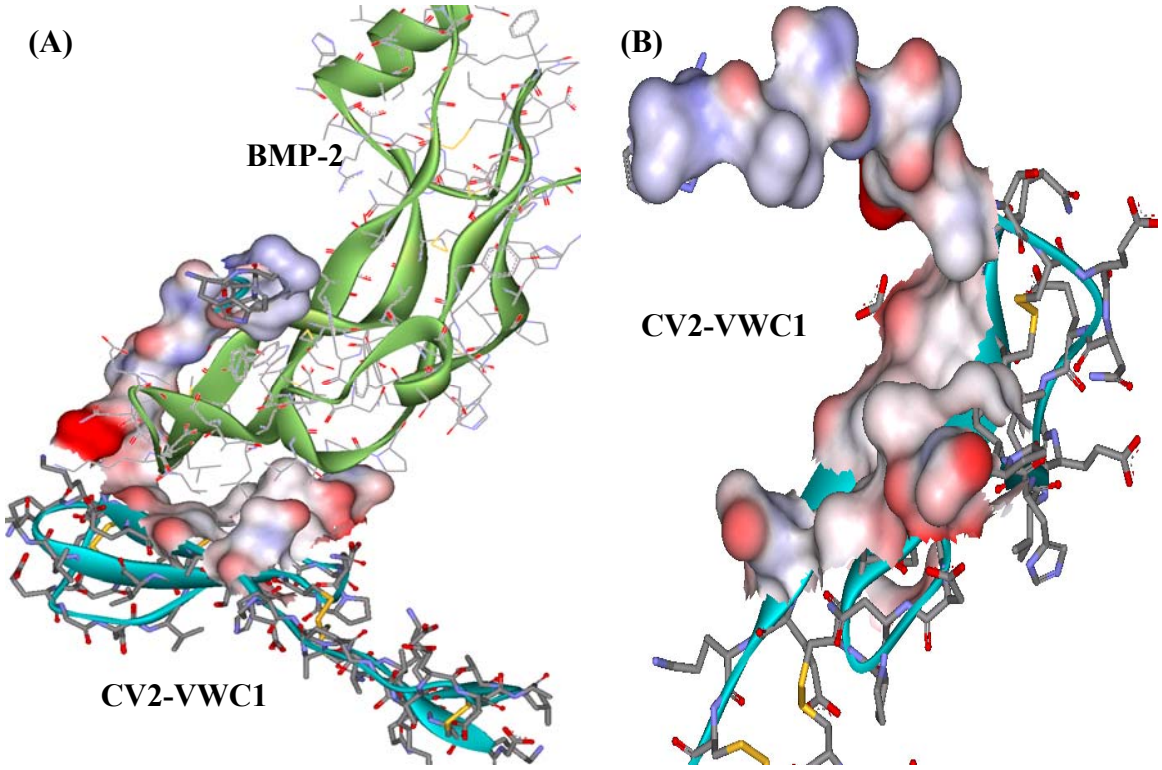

Figure 4.3. The CV2-VWC1 domain surface involved in binding with the BMP-2. (A) Solid surface of the CV2-VWC1 domain residues involved in binding with the BMP-2 dimer. (B) A close snapshot of the CV2-VWC1 binding surface showing several interesting concave groves with hydrophobic and hydrophilic hotspots where small molecules can find tight binding spots.

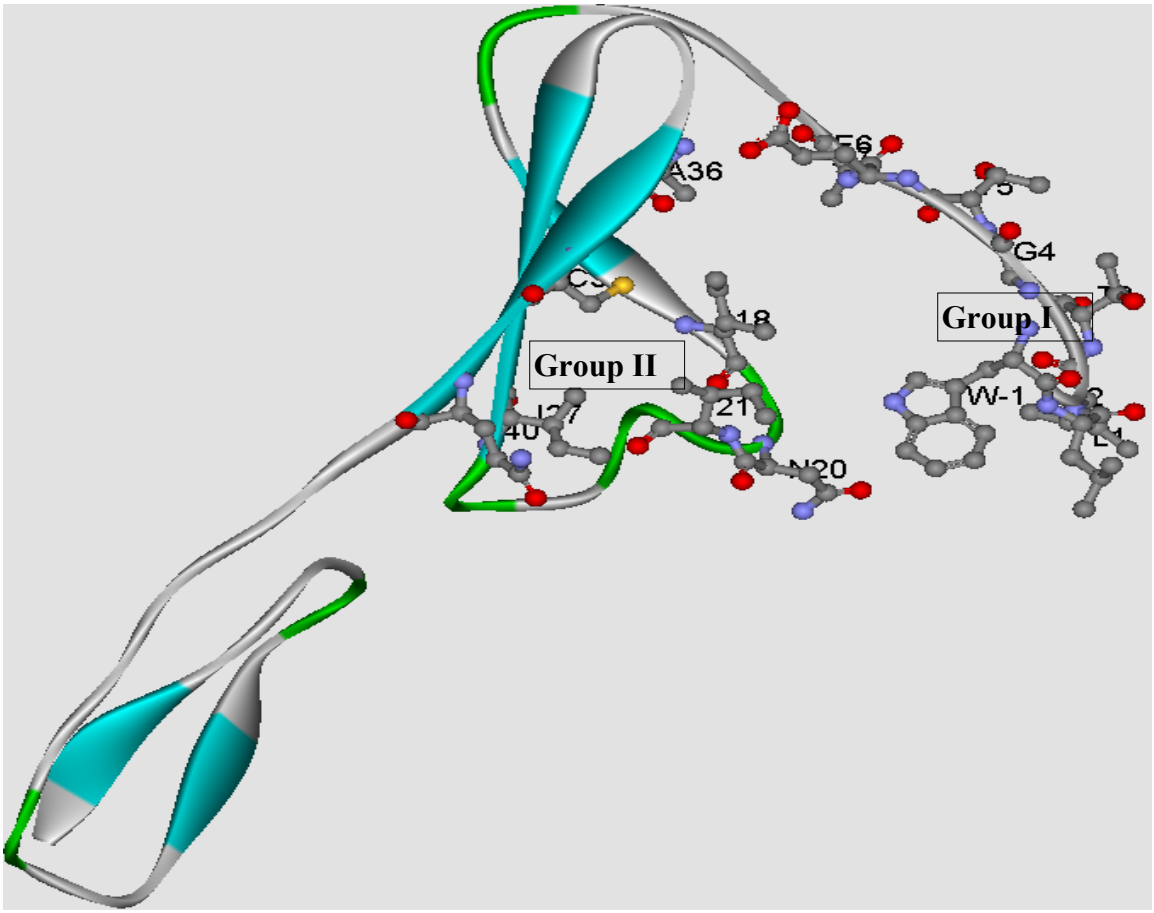

Figure 4.4. The residues of CV2-VWC1 that bind to BMP-2 are divided into two groups as defined for the LUDI runs. The Group I consisting of AAs at positions -1 to 7 with a sphere of 12.0 Å radius and Group II consisting of AAs at positions 18, 20, 21, 27, 36, 38, and 40 with a sphere of 10.0 Å.

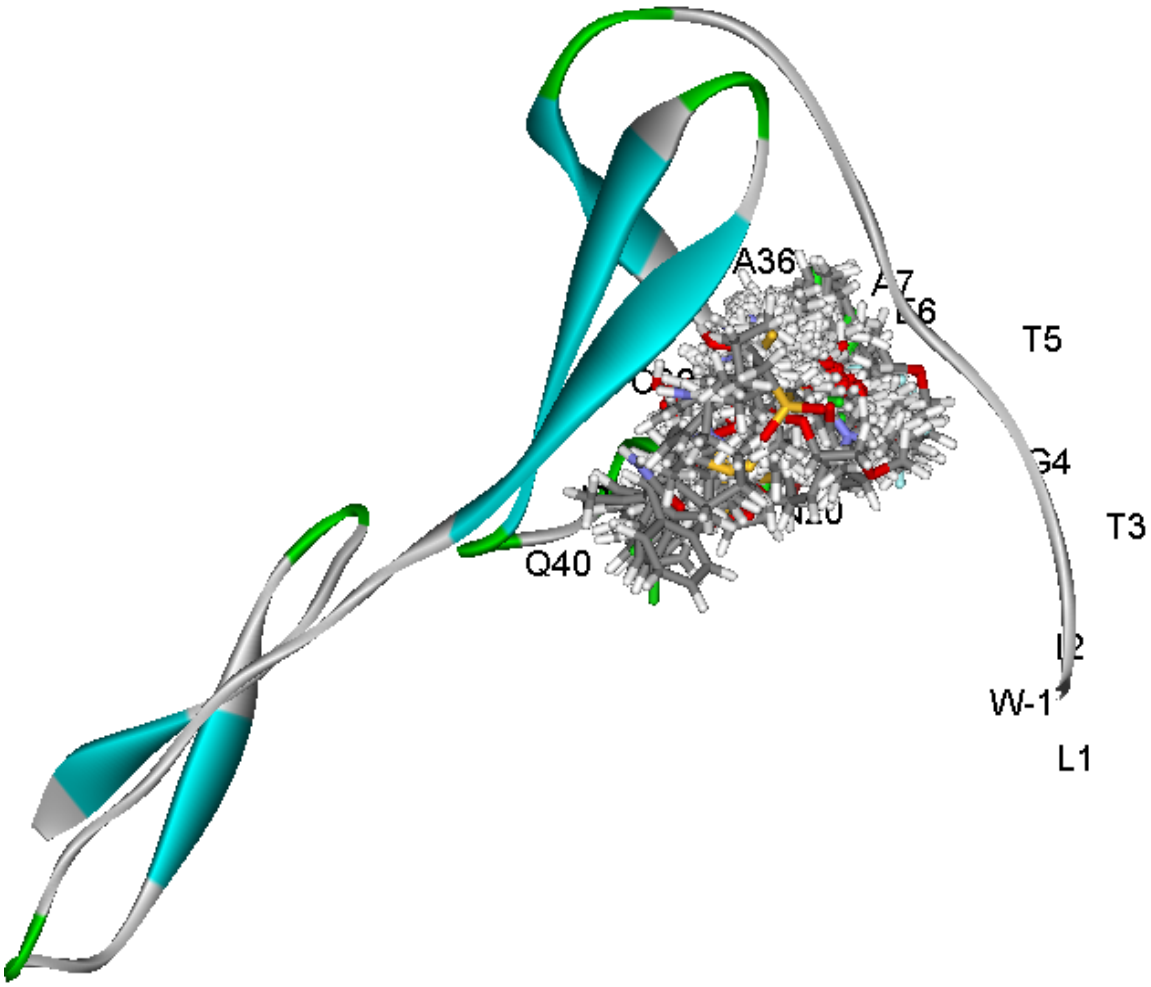

Figure 4.5. Distribution of the top scoring small molecules over the entire binding surface of CV2-VWC1 showing at least two hotspots where most of the top scoring molecules find suitable binding locations.

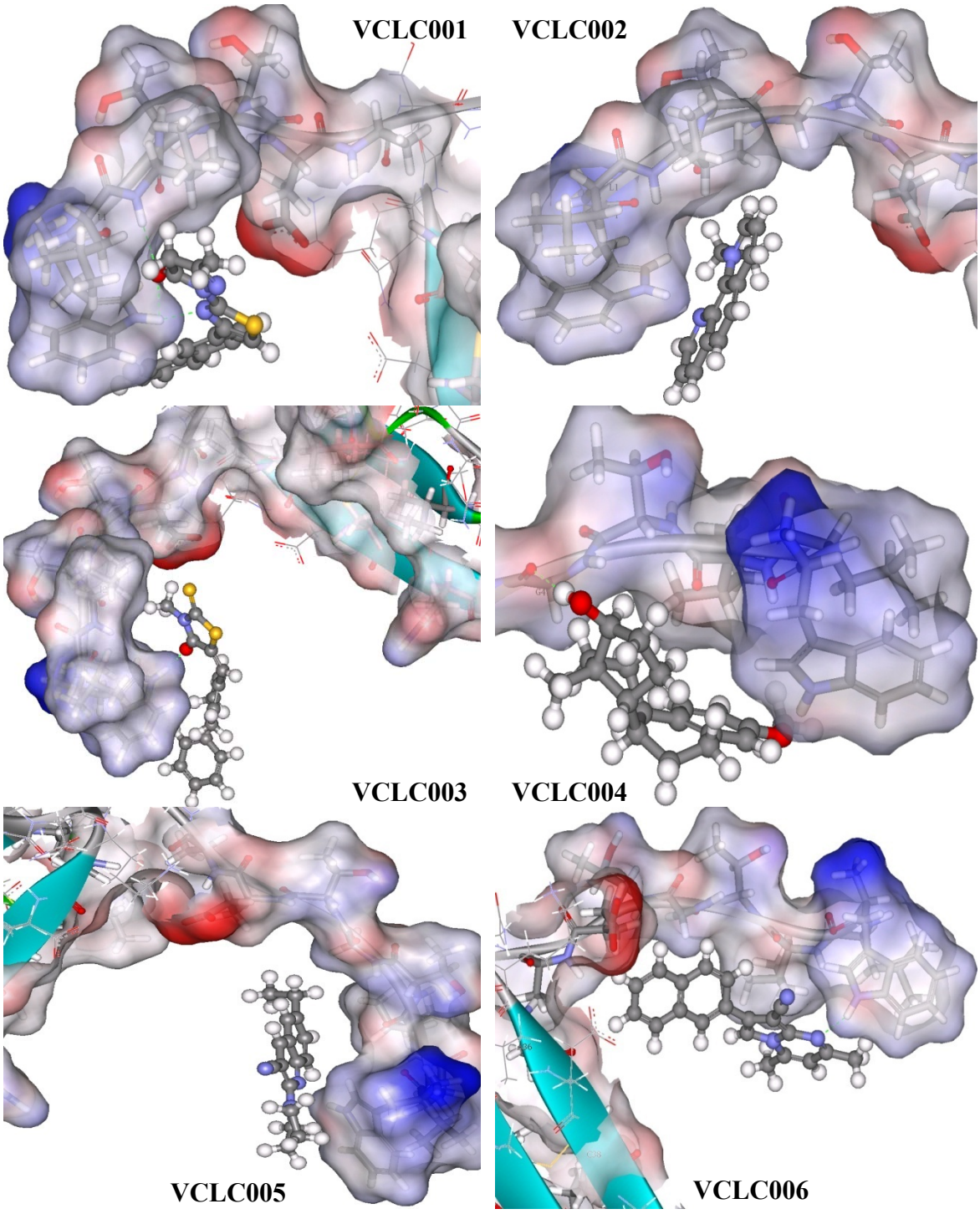

Figure 4.6. Docking poses of some selected top scoring molecules from the CAP database using the LUDI docking procedure.

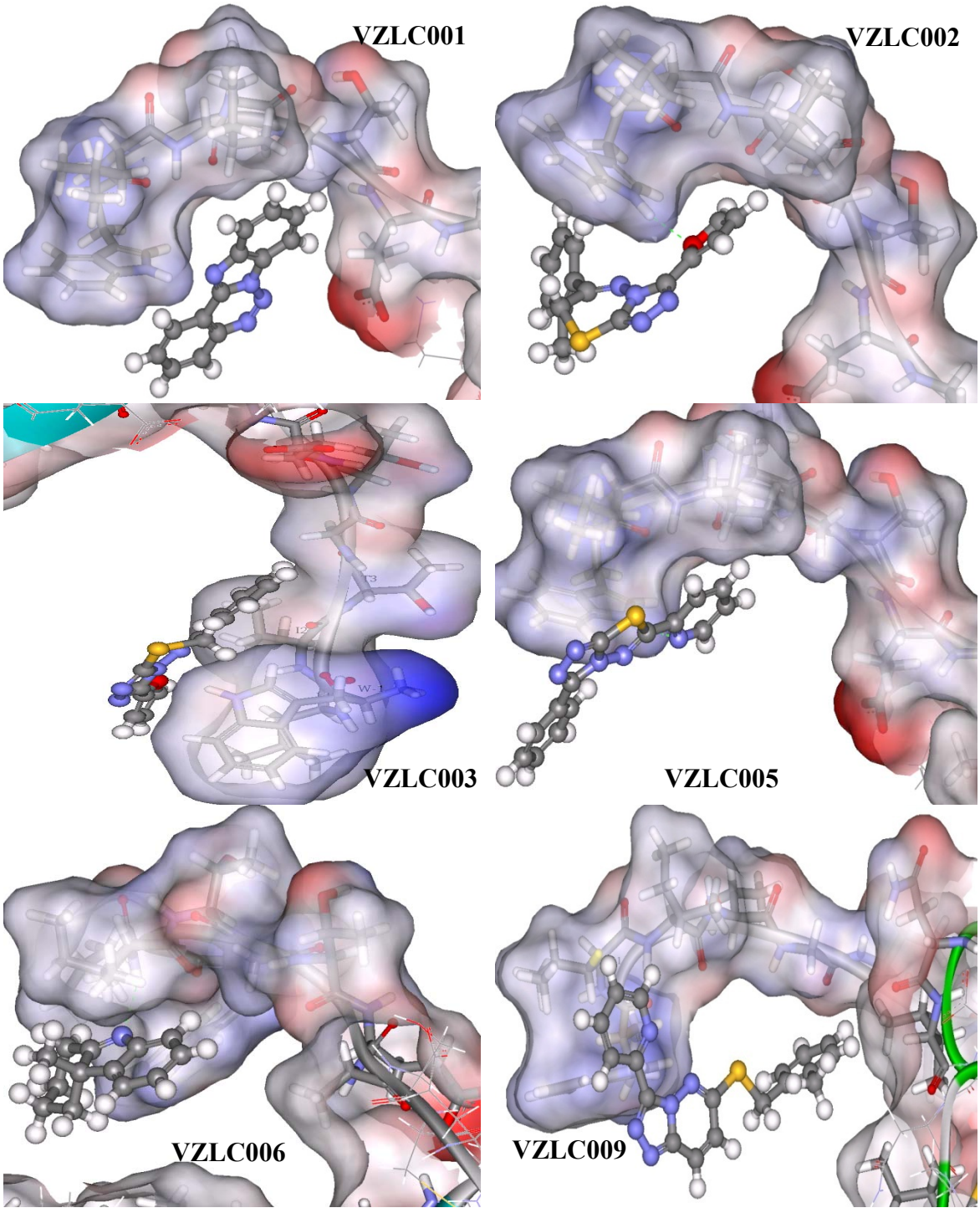

Figure 4.7. Docking poses of some selected top scoring molecules from the ZINC database using the LUDI docking procedure.

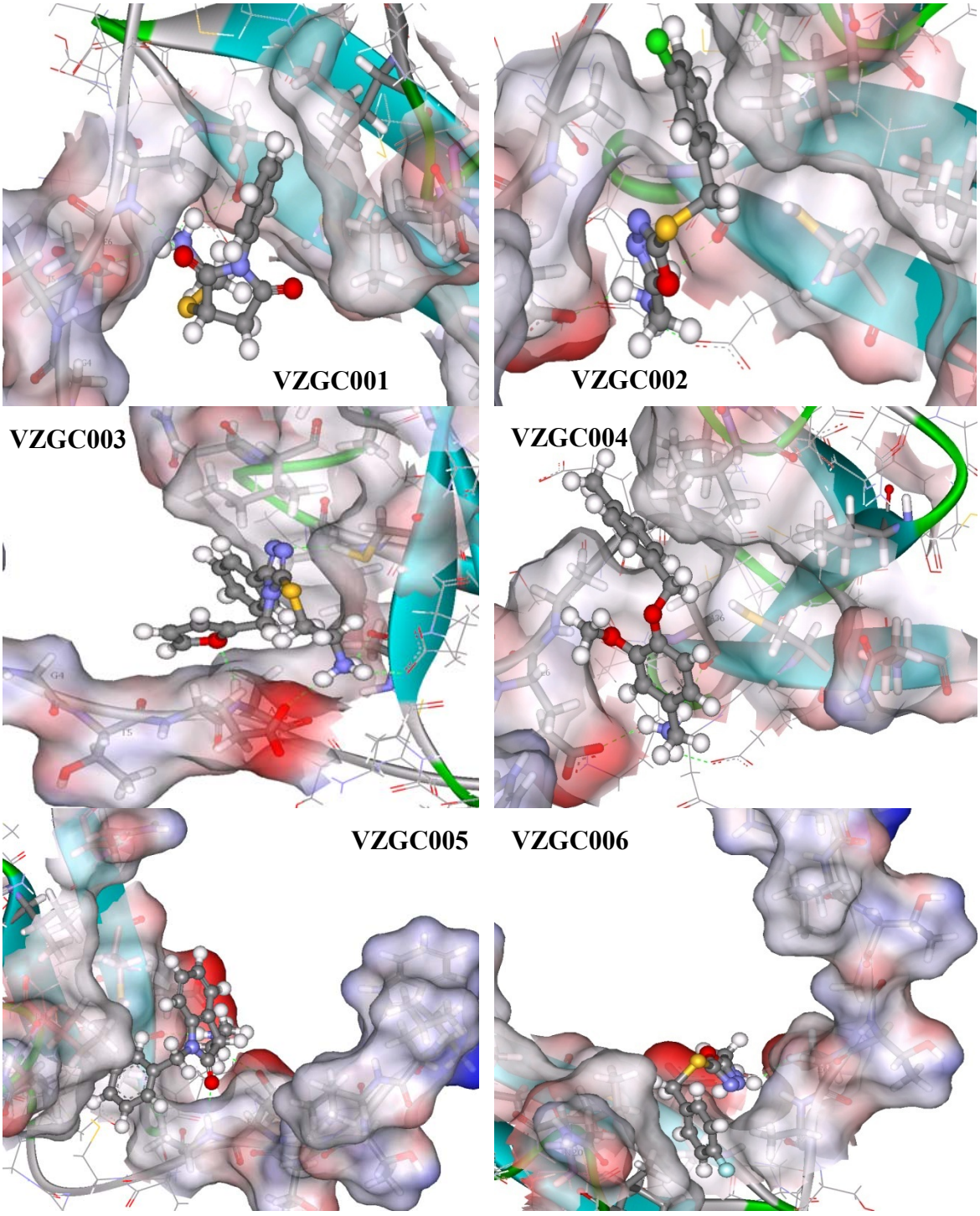

Figure 4.8. Docking poses of some selected top scoring molecules from the ZINC database using the Glide-XP, GOLD, and CombScores.

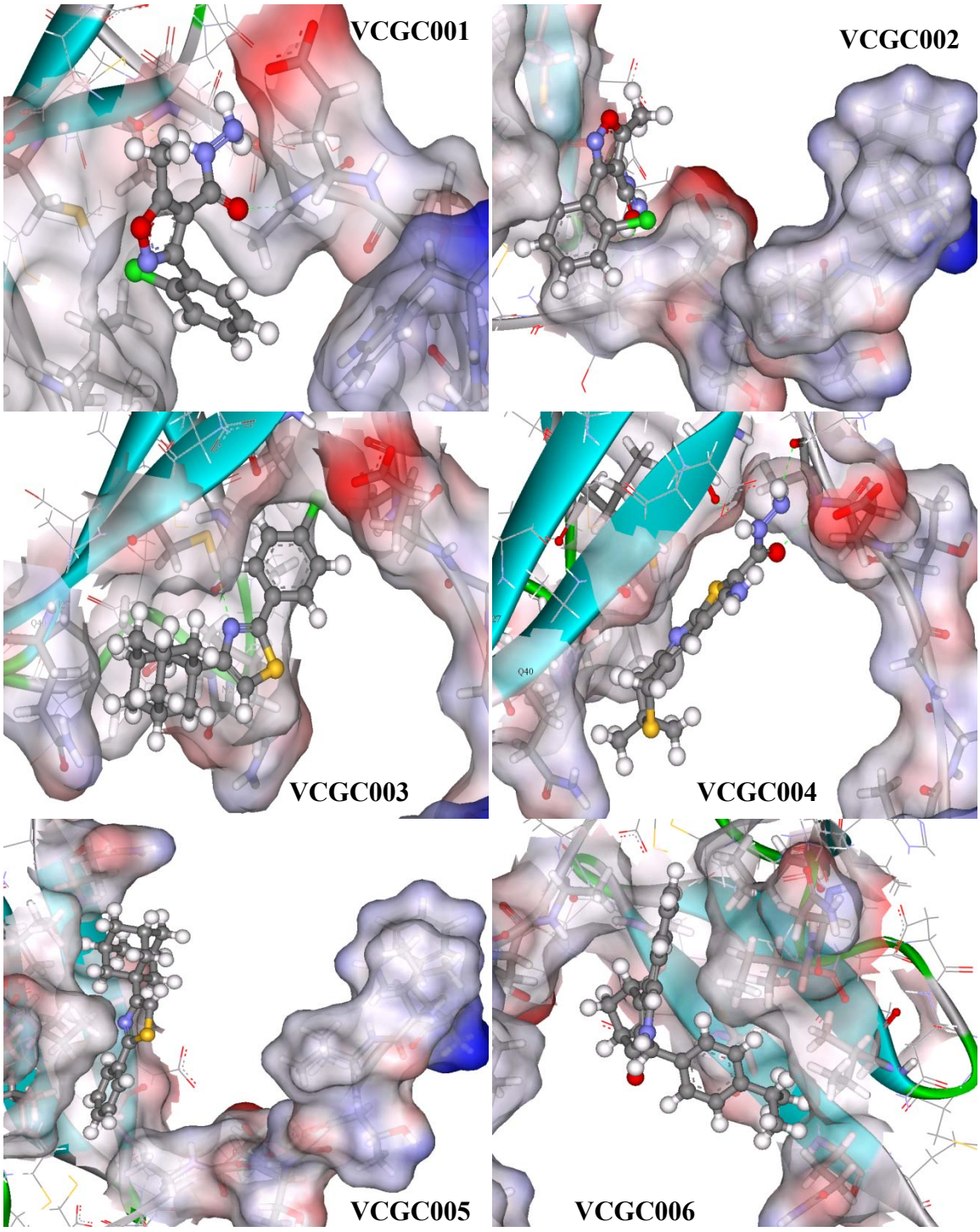

Figure 4.9. Docking poses of some selected top scoring molecules from the CAP database using the Glide, GOLD and CombScores.

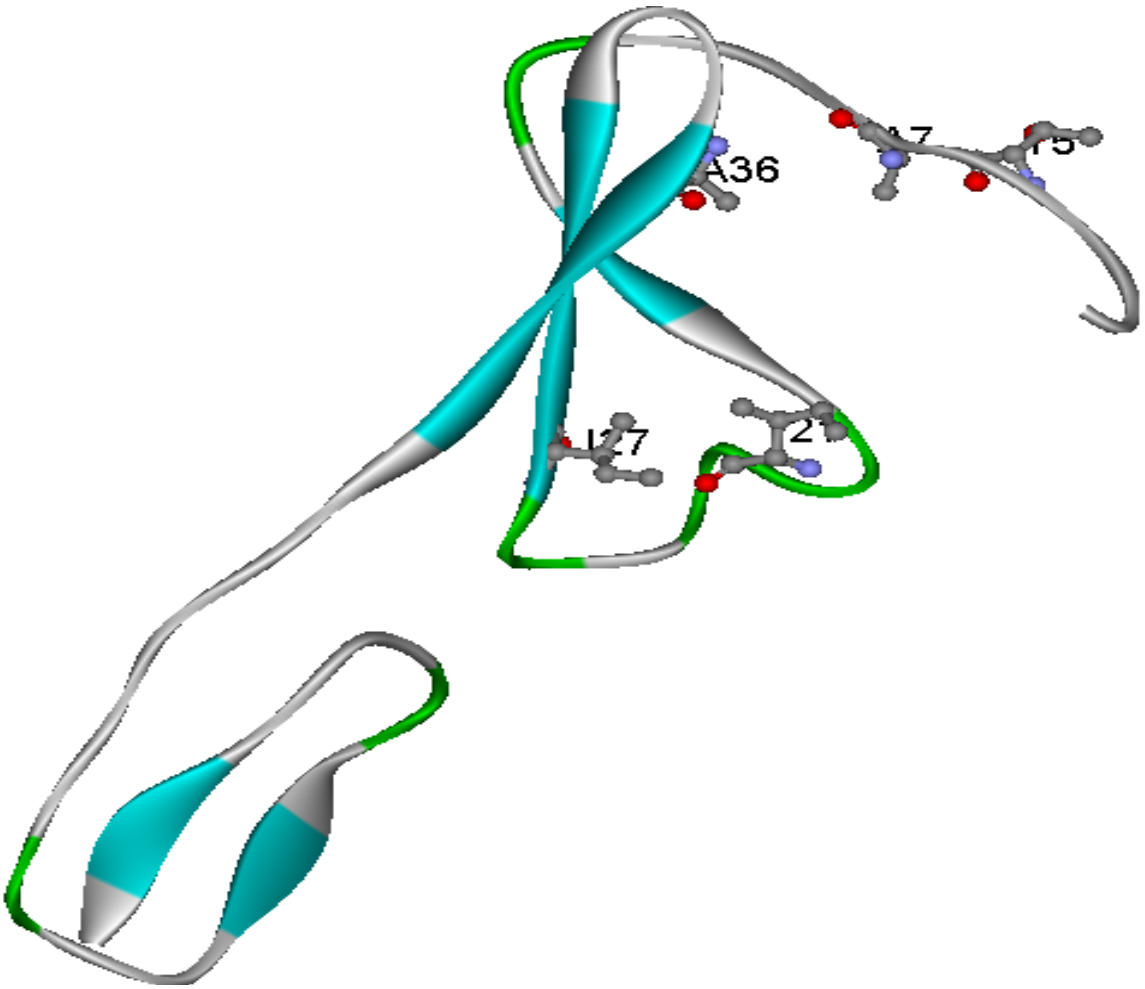

Figure 4.10. Solid ribbon backbone trace of CV2-VWC1 showing some AAs that are in contact with the BMP. Mutational and truncation studies of the AAs, T5, A7, I21, I27, and A36, showed to reduce its binding affinity to BMP, suggesting these could be important target sites. Several of the small molecules in our list shown in tables do block these residues.

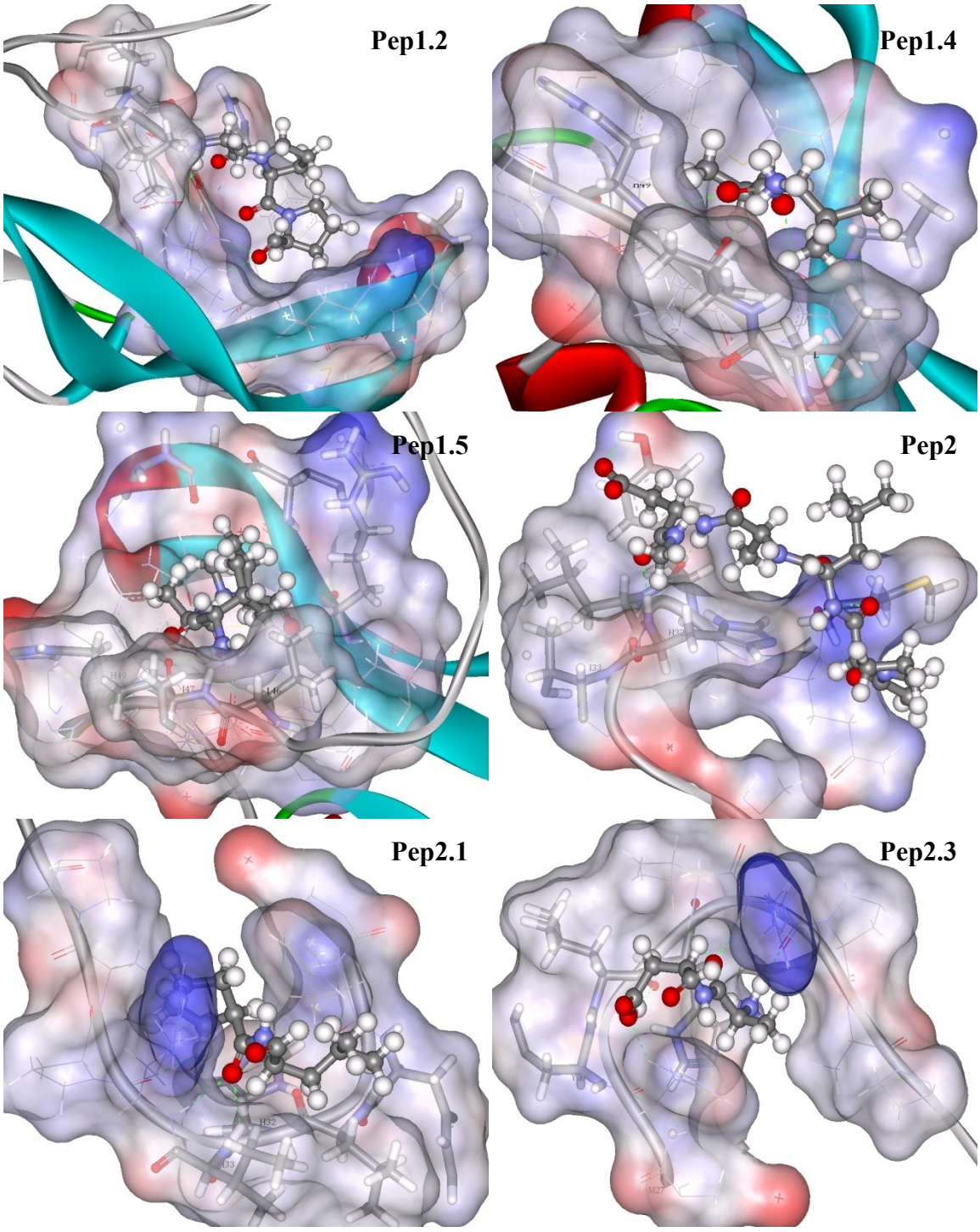

Fig 5.1. Binding poses of some selected representative peptides with high scores from the Glide-XP and the GOLD docking.

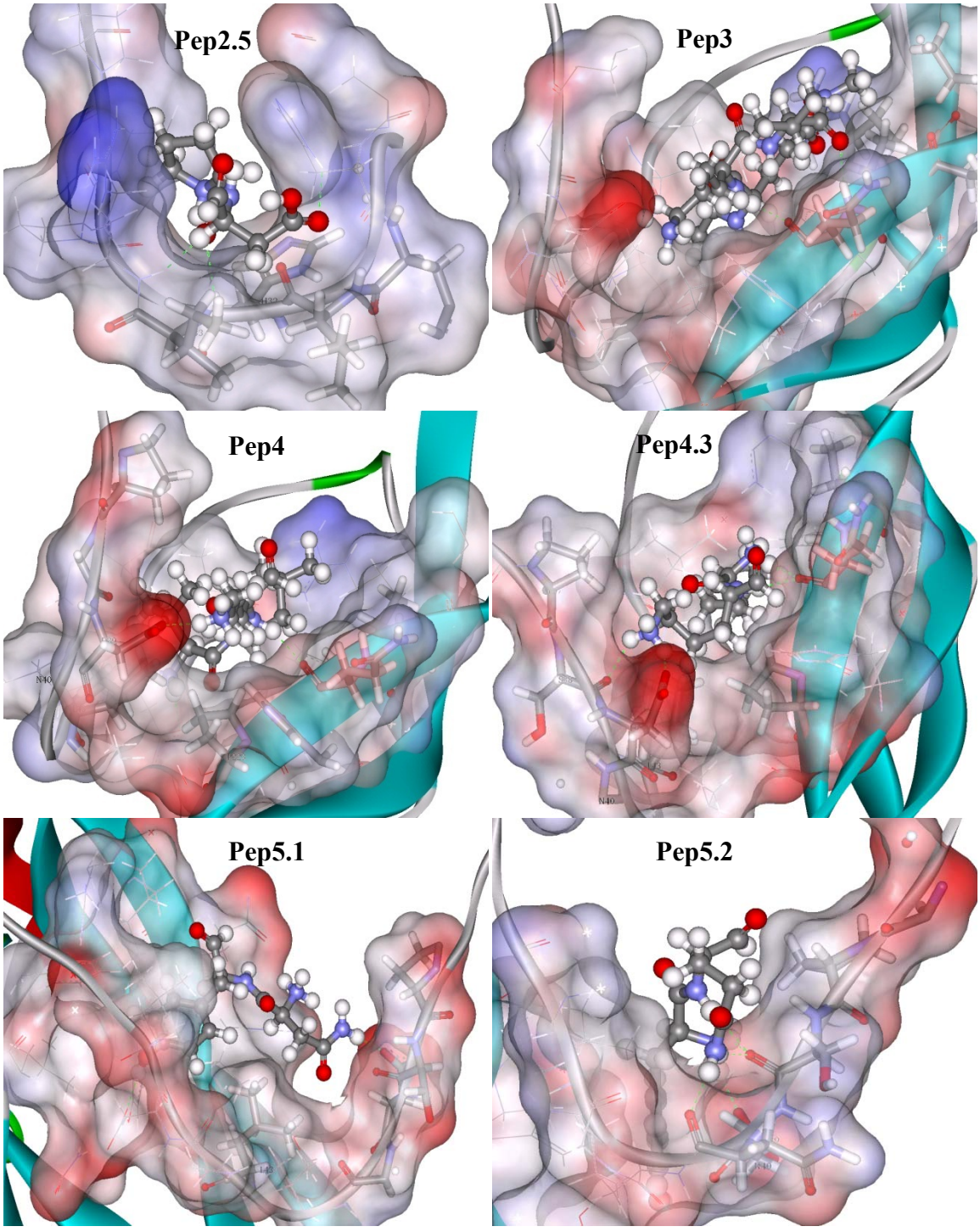

Fig 5.1. (Continued…..)

8.0 RESULTED PUBLICATIONS AND CONFERENCE PRESENTATIONS

8.1 Publications

8.1.1 Manuscripts Accepted /Submitted

- Ahmed S, Metpally RPR, Kumar S, Sangadala S, Reddy, BVB (2009). Computational design of inhibitory agents of BMP/noggin interaction to promote osteogenesis for fast healing of bone fractures. ***International Journal of Biological and Medical Sciences – World Academy of Science, Engineering and Technology 58, 204***
- Ahmed S, Metpally RPR, Sangadala S, Reddy, BVB (2010). Virtual screening and selection of drug-like compounds that block noggin interaction with bone morphogenetic proteins. ***Journal of Molecular Graphics and Modeling*** (in press).
- Ahmed S, Metpally RPR, Sangadala S, Reddy, BVB (2010). In-silico screening for potential inhibitor(s) of follistatin/activin binding to promote activin mediated TGF-β signaling. (Submitted)

8.1.2 Manuscripts in Preparation / Anticipated

- Virtual screening for potential inhibitor(s) of BMP-2/CV2 interaction to promote osteogenesis. (in preparation)
- Modeling and design of peptido-mimetic compounds to block noggin/BMP interaction to promote osteogenesis. (anticipated)

8.2 Meetings/Symposia Presentations

1. Reddy BVB, Sangadala, S, Metpally RPR, **Ahmed S**, Makkar P. Computational design of small molecular inhibitors to promote induction of bone growth by BMPs. At Amerian Chemical Socity (ACS) – 40th Middle Atlantic Regional Meeting (MARM2008) Queens Barrow Community College, New York (May 17-21, 2008) – invited talk by Dr Reddy.

2. **Ahmed S**, Sangadala S, Boden S and Reddy BVB. Computational design of inhibitors of BMP/noggin interaction. At Protein Kinase Targets: Drug Discovery and Design, World Trade Center, Boston, MA (June 4-6, 2007) - Poster

3. **Ahmed S**, Metpally RPR, Sangadala S, Reddy, BVB. Computational design of inhibitors of activin-follistatin interaction. 2nd Advances in Biomolecular Engineering: Protein Design Symposium (June 30, 2008) - The New York. And also at CUNY Graduate Center 9th Annual Celebration of Science Engineering and Mathematics, The Graduate Center, CUNY (April 11, 2008) - Posters.

5. **Ahmed S**, Metpally RPR, Sangadala S, Reddy, BVB. Computational design of inhibitors of BMP/noggin interaction. CUNY Graduate Center: 8th Annual Celebration of Science Engineering and Mathematics on April 13, 2007 - Poster

6. **Ahmed S**, Metpally RPR, Sangadala S, Reddy, BVB. Computational design of small molecular binder to block BMP-Crosveinless 2 interaction. Sigma XI, The Scientific Research Society, Queens College Chapter, Flushing on March 26th, 2009 – Poster.

7. **Ahmed S**, Sangadala S and Reddy BVB. Modeling and Design of Peptido-mimetic Compounds to Block Noggin/BMP, Activin/Follistatin and Crossveinless 2/BMP Interactions to Promote Osteogenesis. 7th Annual Rocky Mountain Bioinformatics Conference, Aspen/Snowmass, Colorado. December 10-12, 2009 – Poster.

8. **Ahmed S**, Sangadala S and Reddy BVB. Modeling and Design of Peptido-mimetic Compounds to Block Noggin/BMP, Activin/Follistatin and Crossveinless 2/BMP Interactions to Promote Osteogenesis. New York Structural Biology Discussion Group - 5th Winter Meeting, The New York Academy of Sciences (January 13th, 2010) 2009 – Poster.

9.0 BIBLIOGRAPHY

1. Vilar JM, Jansen R, Sander C: **Signal processing in the TGF-beta superfamily ligand-receptor network**. *PLoS Comput Biol* 2006, **2**(1):e3.

2. Goumans MJ, Mummery C: **Functional analysis of the TGFbeta receptor/Smad pathway through gene ablation in mice**. *Int J Dev Biol* 2000, **44**(3):253-265.

3. Lin SJ, Lerch TF, Cook RW, Jardetzky TS, Woodruff TK: **The structural basis of TGF-beta, bone morphogenetic protein, and activin ligand binding**. *Reproduction* 2006, **132**(2):179-190.

4. ten Dijke P, Hill CS: **New insights into TGF-beta-Smad signalling**. *Trends Biochem Sci* 2004, **29**(5):265-273.

5. Hartwell LH, Hopfield JJ, Leibler S, Murray AW: **From molecular to modular cell biology**. *Nature* 1999, **402**(6761 Suppl):C47-52.

6. Kitisin K, Saha T, Blake T, Golestaneh N, Deng M, Kim C, Tang Y, Shetty K, Mishra B, Mishra L: **Tgf-Beta signaling in development**. *Sci STKE* 2007, **2007**(399):cm1.

7. Groppe J, Greenwald J, Wiater E, Rodriguez-Leon J, Economides AN, Kwiatkowski W, Baban K, Affolter M, Vale WW, Belmonte JC *et al*: **Structural basis of BMP signaling inhibition by Noggin, a novel twelve-membered cystine knot protein**. *J Bone Joint Surg Am* 2003, **85-A Suppl 3**:52-58.

8. Thompson TB, Lerch TF, Cook RW, Woodruff TK, Jardetzky TS: **The structure of the follistatin:activin complex reveals antagonism of both type I and type II receptor binding**. *Dev Cell* 2005, **9**(4):535-543.

9. Zhang JL, Qiu LY, Kotzsch A, Weidauer S, Patterson L, Hammerschmidt M, Sebald W, Mueller TD: **Crystal structure analysis reveals how the Chordin**

family member crossveinless 2 blocks BMP-2 receptor binding. *Dev Cell* 2008, **14**(5):739-750.

10. Mace PD, Cutfield JF, Cutfield SM: **High resolution structures of the bone morphogenetic protein type II receptor in two crystal forms: implications for ligand binding**. *Biochem Biophys Res Commun* 2006, **351**(4):831-838.

11. Gazzerro E, Minetti C: **Potential drug targets within bone morphogenetic protein signaling pathways**. *Curr Opin Pharmacol* 2007, **7**(3):325-333.

12. Kuo PL, Huang YT, Chang CH, Chang JK: **Bone morphogenetic protein-2 and -4 (BMP-2 and -4) mediates fraxetin-induced maturation and differentiation in human osteoblast-like cell lines**. *Biol Pharm Bull* 2006, **29**(1):119-124.

13. Rosen V, Cox K, Hattersley G: **Bone morphogenetic proteins**. In: *In Principles of Bone Biology***.** 1996: 661-671.

14. Wozney JM: **Bone morphogenetic proteins and their gene expressions**. In: *In Cellular and Molecular Biology of Bone* 1993: 131-167.

15. Sebald W, Nickel J, Zhang JL, Mueller TD: **Molecular recognition in bone morphogenetic protein (BMP)/receptor interaction**. *Biol Chem* 2004, **385**(8):697-710.

16. Rosen V: **BMP and BMP inhibitors in bone**. *Ann N Y Acad Sci* 2006, **1068**:19-25.

17. Lin Y, Martin J, Gruendler C, Farley J, Meng X, Li BY, Lechleider R, Huff C, Kim RH, Grasser WA *et al*: **A novel link between the proteasome pathway and the signal transduction pathway of the bone morphogenetic proteins (BMPs)**. *BMC Cell Biol* 2002, **3**:15.

18. Zimmerman LB, De Jesus-Escobar JM, Harland RM: **The Spemann organizer signal noggin binds and inactivates bone morphogenetic protein 4**. *Cell* 1996, **86**(4):599-606.

19. Re'em-Kalma Y, Lamb T, Frank D: **Competition between noggin and bone morphogenetic protein 4 activities may regulate dorsalization during Xenopus development**. *Proc Natl Acad Sci U S A* 1995, **92**(26):12141-12145.

20. Valenzuela DM, Economides AN, Rojas E, Lamb TM, Nunez L, Jones P, Lp NY, Espinosa R, 3rd, Brannan CI, Gilbert DJ *et al*: **Identification of mammalian noggin and its expression in the adult nervous system**. *J Neurosci* 1995, **15**(9):6077-6084.

21. Larrain J, Bachiller D, Lu B, Agius E, Piccolo S, De Robertis EM: **BMP-binding modules in chordin: a model for signalling regulation in the extracellular space**. *Development* 2000, **127**(4):821-830.

22. Wu XB, Li Y, Schneider A, Yu W, Rajendren G, Iqbal J, Yamamoto M, Alam M, Brunet LJ, Blair HC *et al*: **Impaired osteoblastic differentiation, reduced bone formation, and severe osteoporosis in noggin-overexpressing mice**. *J Clin Invest* 2003, **112**(6):924-934.

23. Groppe J, Greenwald J, Wiater E, Rodriguez-Leon J, Economides AN, Kwiatkowski W, Affolter M, Vale WW, Belmonte JC, Choe S: **Structural basis of BMP signalling inhibition by the cystine knot protein Noggin**. *Nature* 2002, **420**(6916):636-642.

24. Fischer WH, Park M, Donaldson C, Wiater E, Vaughan J, Bilezikjian LM, Vale W: **Residues in the C-terminal region of activin A determine specificity for follistatin and type II receptor binding**. *J Endocrinol* 2003, **176**(1):61-68.

25. Lerch TF, Shimasaki S, Woodruff TK, Jardetzky TS: **Structural and biophysical coupling of heparin and activin binding to follistatin isoform functions**. *J Biol Chem* 2007, **282**(21):15930-15939.

26. Chen YG, Wang Q, Lin SL, Chang CD, Chuang J, Ying SY: **Activin signaling and its role in regulation of cell proliferation, apoptosis, and carcinogenesis**. *Exp Biol Med (Maywood)* 2006, **231**(5):534-544.

27. Vale W, Rivier J, Vaughan J, McClintock R, Corrigan A, Woo W, Karr D, Spiess J: **Purification and characterization of an FSH releasing protein from porcine ovarian follicular fluid**. *Nature* 1986, **321**(6072):776-779.

28. Harrison CA, Gray PC, Koerber SC, Fischer W, Vale W: **Identification of a functional binding site for activin on the type I receptor ALK4**. *J Biol Chem* 2003, **278**(23):21129-21135.

29. Mathews LS, Vale WW: **Expression cloning of an activin receptor, a predicted transmembrane serine kinase**. *Cell* 1991, **65**(6):973-982.

30. Ten Dijke P, Yamashita H, Ichijo H, Franzen P, Laiho M, Miyazono K, Heldin CH: **Characterization of type I receptors for transforming growth factor-beta and activin**. *Science* 1994, **264**(5155):101-104.

31. Wrana JL, Attisano L, Wieser R, Ventura F, Massague J: **Mechanism of activation of the TGF-beta receptor**. *Nature* 1994, **370**(6488):341-347.

32. Wrana JL, Attisano L: **The Smad pathway**. *Cytokine Growth Factor Rev* 2000, **11**(1-2):5-13.

33. Phillips DJ: **Regulation of activin's access to the cell: why is mother nature such a control freak?** *Bioessays* 2000, **22**(8):689-696.

34. Phillips DJ, de Kretser DM: **Follistatin: a multifunctional regulatory protein**. *Front Neuroendocrinol* 1998, **19**(4):287-322.

35. Ueno N, Ling N, Ying SY, Esch F, Shimasaki S, Guillemin R: **Isolation and partial characterization of follistatin: a single-chain Mr 35,000 monomeric protein that inhibits the release of follicle-stimulating hormone**. *Proc Natl Acad Sci U S A* 1987, **84**(23):8282-8286.

36. Nakamura T, Takio K, Eto Y, Shibai H, Titani K, Sugino H: **Activin-binding protein from rat ovary is follistatin**. *Science* 1990, **247**(4944):836-838.

37. Canalis E, Economides AN, Gazzerro E: **Bone morphogenetic proteins, their antagonists, and the skeleton**. *Endocr Rev* 2003, **24**(2):218-235.

38. Chang H, Brown CW, Matzuk MM: **Genetic analysis of the mammalian transforming growth factor-beta superfamily**. *Endocr Rev* 2002, **23**(6):787-823.

39. Matzuk MM, Lu N, Vogel H, Sellheyer K, Roop DR, Bradley A: **Multiple defects and perinatal death in mice deficient in follistatin**. *Nature* 1995, **374**(6520):360-363.

40. Abe Y, Minegishi T, Leung PC: **Activin receptor signaling**. *Growth Factors* 2004, **22**(2):105-110.

41. Gumienny TL, Padgett RW: **The other side of TGF-beta superfamily signal regulation: thinking outside the cell**. *Trends Endocrinol Metab* 2002, **13**(7):295-299.

42. Schneyer AL, Wang Q, Sidis Y, Sluss PM: **Differential distribution of follistatin isoforms: application of a new FS315-specific immunoassay**. *J Clin Endocrinol Metab* 2004, **89**(10):5067-5075.

43. Schneyer AL, Hall HA, Lambert-Messerlian G, Wang QF, Sluss P, Crowley WF, Jr.: **Follistatin-activin complexes in human serum and follicular fluid differ immunologically and biochemically**. *Endocrinology* 1996, **137**(1):240-247.

44. Hashimoto O, Nakamura T, Shoji H, Shimasaki S, Hayashi Y, Sugino H: **A novel role of follistatin, an activin-binding protein, in the inhibition of activin action in rat pituitary cells. Endocytotic degradation of activin and its acceleration by follistatin associated with cell-surface heparan sulfate**. *J Biol Chem* 1997, **272**(21):13835-13842.

45. Wang Q, Keutmann HT, Schneyer AL, Sluss PM: **Analysis of human follistatin structure: identification of two discontinuous N-terminal sequences coding for activin A binding and structural consequences of activin binding to native proteins**. *Endocrinology* 2000, **141**(9):3183-3193.

46. Keutmann HT, Schneyer AL, Sidis Y: **The role of follistatin domains in follistatin biological action**. *Mol Endocrinol* 2004, **18**(1):228-240.

47. Harrison CA, Chan KL, Robertson DM: **Activin-A binds follistatin and type II receptors through overlapping binding sites: generation of mutants with isolated binding activities**. *Endocrinology* 2006, **147**(6):2744-2753.

48. Shimasaki S, Koga M, Buscaglia ML, Simmons DM, Bicsak TA, Ling N: **Follistatin gene expression in the ovary and extragonadal tissues**. *Mol Endocrinol* 1989, **3**(4):651-659.

49. Ullman CG, Perkins SJ: **The Factor I and follistatin domain families: the return of a prodigal son**. *Biochem J* 1997, **326 (Pt 3)**:939-941.

50. Sugino K, Kurosawa N, Nakamura T, Takio K, Shimasaki S, Ling N, Titani K, Sugino H: **Molecular heterogeneity of follistatin, an activin-binding protein. Higher affinity of the carboxyl-terminal truncated forms for heparan sulfate proteoglycans on the ovarian granulosa cell**. *J Biol Chem* 1993, **268**(21):15579-15587.

51. Harrington AE, Morris-Triggs SA, Ruotolo BT, Robinson CV, Ohnuma S, Hyvonen M: **Structural basis for the inhibition of activin signalling by follistatin**. *Embo J* 2006, **25**(5):1035-1045.

52. Inouye S, Ling N, Shimasaki S: **Localization of the heparin binding site of follistatin**. *Mol Cell Endocrinol* 1992, **90**(1):1-6.

53. Innis CA, Hyvonen M: **Crystal structures of the heparan sulfate-binding domain of follistatin. Insights into ligand binding**. *J Biol Chem* 2003, **278**(41):39969-39977.

54. Shimonaka M, Inouye S, Shimasaki S, Ling N: **Follistatin binds to both activin and inhibin through the common subunit**. *Endocrinology* 1991, **128**(6):3313-3315.

55. Thompson TB, Woodruff TK, Jardetzky TS: **Structures of an ActRIIB:activin A complex reveal a novel binding mode for TGF-beta ligand:receptor interactions**. *Embo J* 2003, **22**(7):1555-1566.

56. Harrison CA, Gray PC, Fischer WH, Donaldson C, Choe S, Vale W: **An activin mutant with disrupted ALK4 binding blocks signaling via type II receptors**. *J Biol Chem* 2004, **279**(27):28036-28044.

57. Kirsch T, Sebald W, Dreyer MK: **Crystal structure of the BMP-2-BRIA ectodomain complex**. *Nat Struct Biol* 2000, **7**(6):492-496.

58. Hogan BL: **Bone morphogenetic proteins: multifunctional regulators of vertebrate development**. *Genes Dev* 1996, **10**(13):1580-1594.

59. Massague J, Chen YG: **Controlling TGF-beta signaling**. *Genes Dev* 2000, **14**(6):627-644.

60. Little SC, Mullins MC: **Extracellular modulation of BMP activity in patterning the dorsoventral axis**. *Birth Defects Res C Embryo Today* 2006, **78**(3):224-242.

61. Ikeya M, Kawada M, Kiyonari H, Sasai N, Nakao K, Furuta Y, Sasai Y: **Essential pro-Bmp roles of crossveinless 2 in mouse organogenesis**. *Development* 2006, **133**(22):4463-4473.

62. Zakin L, Metzinger CA, Chang EY, Coffinier C, De Robertis EM: **Development of the vertebral morphogenetic field in the mouse: interactions between Crossveinless-2 and Twisted Gastrulation**. *Dev Biol* 2008, **323**(1):6-18.

63. Rentzsch F, Zhang J, Kramer C, Sebald W, Hammerschmidt M: **Crossveinless 2 is an essential positive feedback regulator of Bmp signaling during zebrafish gastrulation**. *Development* 2006, **133**(5):801-811.

64. Binnerts ME, Wen X, Cante-Barrett K, Bright J, Chen HT, Asundi V, Sattari P, Tang T, Boyle B, Funk W *et al*: **Human Crossveinless-2 is a novel inhibitor of**

bone morphogenetic proteins. *Biochem Biophys Res Commun* 2004, **315**(2):272-280.

65. Moser M, Binder O, Wu Y, Aitsebaomo J, Ren R, Bode C, Bautch VL, Conlon FL, Patterson C: **BMPER, a novel endothelial cell precursor-derived protein, antagonizes bone morphogenetic protein signaling and endothelial cell differentiation**. *Mol Cell Biol* 2003, **23**(16):5664-5679.

66. Kamimura M, Matsumoto K, Koshiba-Takeuchi K, Ogura T: **Vertebrate crossveinless 2 is secreted and acts as an extracellular modulator of the BMP signaling cascade**. *Dev Dyn* 2004, **230**(3):434-445.

67. Ambrosio AL, Taelman VF, Lee HX, Metzinger CA, Coffinier C, De Robertis EM: **Crossveinless-2 Is a BMP feedback inhibitor that binds Chordin/BMP to regulate Xenopus embryonic patterning**. *Dev Cell* 2008, **15**(2):248-260.

68. Gurdon JB, Bourillot PY: **Morphogen gradient interpretation**. *Nature* 2001, **413**(6858):797-803.

69. Serpe M, Umulis D, Ralston A, Chen J, Olson DJ, Avanesov A, Othmer H, O'Connor MB, Blair SS: **The BMP-binding protein Crossveinless 2 is a short-range, concentration-dependent, biphasic modulator of BMP signaling in Drosophila**. *Dev Cell* 2008, **14**(6):940-953.

70. Conley CA, Silburn R, Singer MA, Ralston A, Rohwer-Nutter D, Olson DJ, Gelbart W, Blair SS: **Crossveinless 2 contains cysteine-rich domains and is required for high levels of BMP-like activity during the formation of the cross veins in Drosophila**. *Development* 2000, **127**(18):3947-3959.

71. Zhang JL, Huang Y, Qiu LY, Nickel J, Sebald W: **von Willebrand factor type C domain-containing proteins regulate bone morphogenetic protein signaling through different recognition mechanisms**. *J Biol Chem* 2007, **282**(27):20002-20014.

72. Qiu LY, Zhang JL, Kotzsch A, Sebald W, Mueller TD: **Crystallization and preliminary X-ray analysis of the complex of the first von Willebrand type C domain bound to bone morphogenetic protein 2**. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 2008, **64**(Pt 4):307-312.

73. Abreu JG, Ketpura NI, Reversade B, De Robertis EM: **Connective-tissue growth factor (CTGF) modulates cell signalling by BMP and TGF-beta**. *Nat Cell Biol* 2002, **4**(8):599-604.

74. Garcia Abreu J, Coffinier C, Larrain J, Oelgeschlager M, De Robertis EM: **Chordin-like CR domains and the regulation of evolutionarily conserved extracellular signaling systems**. *Gene* 2002, **287**(1-2):39-47.

75. Fairlie DP, West ML, Wong AK: **Towards protein surface mimetics**. *Curr Med Chem* 1998, **5**(1):29-62.

76. Shrivastava A, Nunn AD, Tweedle MF: **Designer peptides: learning from nature**. *Curr Pharm Des* 2009, **15**(6):675-681.

77. Schrader T, Koch S: **Artificial protein sensors**. *Mol Biosyst* 2007, **3**(4):241-248.

78. Robinson JA: **Beta-hairpin peptidomimetics: design, structures and biological activities**. *Acc Chem Res* 2008, **41**(10):1278-1288.

79. Gasteiger E, Jung E, Bairoch A: **SWISS-PROT: connecting biomolecular knowledge via a protein database**. *Curr Issues Mol Biol* 2001, **3**(3):47-55.

80. Irwin JJ, Shoichet BK: **ZINC--a free database of commercially available compounds for virtual screening**. *J Chem Inf Model* 2005, **45**(1):177-182.

81. Thompson JD, Higgins DG, Gibson TJ: **CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice**. *Nucleic Acids Res* 1994, **22**(22):4673-4680.

82. Richmond TJ, Richards FM: **Packing of alpha-helices: geometrical constraints and contact areas**. *J Mol Biol* 1978, **119**(4):537-555.

83. Bohm HJ: **The computer program LUDI: a new method for the de novo design of enzyme inhibitors**. *J Comput Aided Mol Des* 1992, **6**(1):61-78.

84. Bohm HJ: **On the use of LUDI to search the Fine Chemicals Directory for ligands of proteins of known three-dimensional structure**. *J Comput Aided Mol Des* 1994, **8**(5):623-632.

85. Halgren TA, Murphy RB, Friesner RA, Beard HS, Frye LL, Pollard WT, Banks JL: **Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening**. *J Med Chem* 2004, **47**(7):1750-1759.

86. Kirkpatrick P: **Gliding to success**. *Nature Reviews Drug Discovery* 2004, **3**:299.

87. Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, Repasky MP, Knoll EH, Shelley M, Perry JK *et al*: **Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy**. *J Med Chem* 2004, **47**(7):1739-1749.

88. Verdonk ML, Chessari G, Cole JC, Hartshorn MJ, Murray CW, Nissink JW, Taylor RD, Taylor R: **Modeling water molecules in protein-ligand docking using GOLD**. *J Med Chem* 2005, **48**(20):6504-6515.

89. Verdonk ML, Cole JC, Hartshorn MJ, Murray CW, Taylor RD: **Improved protein-ligand docking using GOLD**. *Proteins* 2003, **52**(4):609-623.

90. Verdonk ML, Berdini V, Hartshorn MJ, Mooij WT, Murray CW, Taylor RD, Watson P: **Virtual screening using protein-ligand docking: avoiding artificial enrichment**. *J Chem Inf Comput Sci* 2004, **44**(3):793-806.

91. Banks JL, Beard HS, Cao Y, Cho AE, Damm W, Farid R, Felts AK, Halgren TA, Mainz DT, Maple JR *et al*: **Integrated Modeling Program, Applied Chemical Theory (IMPACT)**. *J Comput Chem* 2005, **26**(16):1752-1780.

92. Maiti R, Van Domselaar GH, Zhang H, Wishart DS: **SuperPose: a simple server for sophisticated structural superposition**. *Nucleic Acids Res* 2004, **32**(Web Server issue):W590-594.

93. Kontoyianni M, McClellan LM, Sokol GS: **Evaluation of docking performance: comparative data on docking algorithms**. *J Med Chem* 2004, **47**(3):558-565.

94. Perola E, Walters WP, Charifson PS: **A detailed comparison of current docking and scoring methods on systems of pharmaceutical relevance**. *Proteins* 2004, **56**(2):235-249.

95. Nagamine N, Shirakawa T, Minato Y, Torii K, Kobayashi H, Imoto M, Sakakibara Y: **Integrating statistical predictions and experimental verifications for enhancing protein-chemical interaction predictions in virtual screening**. *PLoS Comput Biol* 2009, **5**(6):e1000397.

96. Rastelli G, Rio AD, Degliesposti G, Sgobba M: **Fast and accurate predictions of binding free energies using MM-PBSA and MM-GBSA**. *J Comput Chem* 2009.