

NEW CONFORMATIONAL PROPERTIES OF SH2 DOMAIN BINDING POCKET

V. V. HURMACH¹, M. O. PLATONOV², Yu. I. PRYLUTSKYY¹

¹Taras Shevchenko National University of Kyiv, Ukraine;

²Institute of Molecular Biology and Genetics,
National Academy of Sciences of Ukraine, Kyiv;
e-mail: gyrmach@gmail.com

The conformational changes of proteins play an important role in biological functioning such as ligand-protein and protein-protein interactions. The aim of the work was to investigate the conformational movement of most represented SH2 domains. It was found that SH2 domain binding pocket includes both flexible and not flexible regions: the central area of the binding pocket is the most unflexible, whereas the pTyr-binding and hydrophobic zones are the most flexible. Results of the computer analysis revealed new conformational properties of SH2 domain, which are important for drug design.

Key words: SH2 domain, conformational movement, binding pocket.

It is known that proteins undergo conformational changes in the course of their functioning. Conformations, responsible for a particular property, may not correspond to the global minimum, and in some cases the ensemble of conformations can be responsible for an observed behavior. So, the conformation with minimal energy is not enough for representing the molecule. Conformational changes may range from small local movements to large domain motions, which affect the protein functions. Typical example is the transformation of enzyme active site (close, open) that provides chemical modification of ligand (increasing selectivity to active site of enzyme). That is why conformational analysis of protein plays an important role in the field of bioinformatics for protein secondary structure prediction and drug design.

It is known that Src homology 2 (SH2) domains are small protein modules consisting of approximately 100 amino acids which form two α -helix and seven β -sheets. Human genome encodes about 120 SH2 domains which are included in 110 different proteins, such as protein kinases (Src, Lck), phosphatases (SHP2, SHIP2), phospholipases (PLC1), transcription factors (STAT), regulatory proteins (SOCS), adapter proteins (Grb2), structural proteins (SHC) and others. These domains play an important role in intracellular signaling in complex with different protein, mainly Src oncoprotein [1]. Binding pocket of SH2 domain is characterized by a well-

defined placement of pTyr-binding part (usually ArgB5 creates a strong hydrogen bond with phosphotyrosine), a central part (hydrogen bond between NH group of ligand and O of His) and a large hydrophobic part (interacts with a hydrophobic part of ligand) [2-4]. The main function of SH2 domain is pTyr-binding recognition [5]. Recent research has shown that each SH2 domain binds only to specific phosphotyrosine-containing fragments (e.g. the Src SH2 domain recognizes GluGlu-Ile (binding fragment pYEEI), Grb2 SH2 domain binds to pYVNV) [6]. Based on this, it is possible to classify SH2 domains according to the specificity of recognition of pTyr residue with C-terminus [6].

In the work we analyzed different possible conformations of SH2 domains. For generation of the conformations the LowModeMD Search method was used. It employs a short ~ 1 ps run of Molecular Dynamics (MD) at constant temperature followed by all-atom energy minimization [7].

The aim of the work was to define the main flexible and rigid regions of SH2 domain. However, we must note that complete understanding of this effect requires detailed study of thermodynamic properties of SH2 domains.

Materials and Methods

Eight SH2 domains were used for conformational search, namely: 2KK6 (solution structure of SH2 domain of proto-oncogene tyrosine-protein ki-

nase, *Homo sapiens*), 3IN7 (crystal structure of the Grb2 SH2 domain in complex with a Cyclopropyl-constrained Ac-pY-Q-N-NH₂ Tripeptide Mimic, *Homo sapiens*), 2K7A (complex between the SH3 and SH2 domain of interleukin-2 tyrosine kinase, *Mus musculus*), 2JYQ (NMR structure of the apo v-Src SH2 domain, Rous sarcoma virus), 2GE9 (solution structures of the SH2 domain of Bruton's Tyrosine Kinase, *Homo sapiens*), 2FCI (Syk tyrosine kinase, *Bos taurus*), 1UUS (structure of an activated D. STAT in its DNA-unbound form, *Dictyostelium discoideum*) and 1O49 (crystal structure of SH2 proto-oncogene tyrosine protein kinase SRC, *Homo sapiens*).

Search for new conformations of SH2 domains was performed by using MOE software package [8] where Low Mode Search (a short molecular dynamics simulation using velocities with small kinetic energy on the high-frequency vibrational modes) method was applied. So, the conformation of SH2 domain was generated according to the following options:

- Rejection Limit (the numbers of contiguous attempts which must be made to generate new conformation prior to terminating the search) – 100;

- Iteration Limit (maximum number of attempts to generate new conformation (independent of the uniqueness of each generated conformation)) – 10;

- RMS Gradient (if root mean square gradient falls below the specified value the energy minimization is terminated) – 0.005;

- MM Iteration Limit (the maximum number of energy minimization steps performed for each conformer energy minimization) – 50;

- RMSD Limit (if two conformations have a Rmsd less than the specified value, they are considered indistinguishable) – 0.25 Å, 0.5 Å and 0.75 Å;

- Strain Cutoff (energy minimized conformations will be discarded if their potential energy is greater than $E_{\min} + s$, where E_{\min} is the lowest energy among the minimized conformations and s is the value specified for Strain Cutoff (in kcal/mol)) – 7.

The obtained conformations were energy minimized in Cartesian coordinates system.

Result and Discussion

The mobilities of eight most representative SH2 domains were determined. The example of SH2 domain conformation superposition is presented in Fig. 1. RMSD (root mean square deviations) for all obtained conformations are given in Table and for

all residues are presented in Fig 2. Generally, 39 (from four to six per structure) conformations were obtained. The observed average RMSD of any conformation varies between 1 Å and 3.5 Å, and only in one case it is equal to 5.17 Å (conformation N3 of PDB ID 1O49). To take into consideration that “DE-loop, β E – sheet, EF-loop and FB-loop” region is located only in one part of domain, we examined them together in Fig. 1.

Movement of the secondary structure of SH2 domain. In general, SH2 domain consists of a large number of secondary structures (AA-loop, AB-loop, BC-loop, CD-loop, BG-loop, EF-loop, DE-loop, BG-loop, α A-helix, α B-helix, β A-sheet, β B-sheet, β C-sheet, β D-sheet, β E-sheet, β F-sheet, β G-sheet) with different flexibility. Obtained data demonstrate essential contribution of α A-helix (average RMSD range for the obtained conformations: 1.5-2.7 Å), BC-loop (1.3-3.1 Å), CD-loop (1.3-6.7 Å), α B-helix (1.0-2.5 Å), BG-loop (1.8-2.5 Å), DE-loop, β E-sheet, EF-loop and FB-loop (1.6-2.8 Å)) secondary structures to the total variability of SH2 domain conformations.

Nevertheless, the mobility of different parts of the above mentioned secondary structures is not equal. The α A-helix (1O49 – RRESERLL, 3IN7 – RAKAEEML, 1UUS – RQEVNDA, 2FCI – RAQAEHMLMR, 2GE9 – RSQAEQLLKQE, 2JYQ – RRESERLL, 2K7A – RDKAEKLLD, 2KK6 – RIEAQE) contains two high flexible amino acid (number one (0.4-7.5 Å) and five (0.724-5.140 Å), e. g. in PDB 1O49: RRESERLL – in red). In contrast, in all conformations of 2KK6 the most flexible is the amino acid number 2 (2.5-5.7 Å). All other amino acids of the helix are less flexible (the range of flexibility: 1.0-2.5 Å). The similar situation is in the BG-loop (1O49 – HADGLCHRLTVCPT, 1UUS – HSQFV, 2FCI – HPLYRKMK, 2GE9 – HNSAGLISRLKYPVSQQNKNAPST, 2K7A – YNGGGLVTRLRYPVCG, 2JYQ – SKHADGLCHRLTNVC, 2KK6 – TKQVITKSGVLLNPIP, 3IN7 – STSVSRNQQIFLRDIEQ). The average conformational flexibility of 1O49, 2FCI, 2GE9, 2K7A, 2JYQ, 2JYQ and 2KK6 is in the range of 2.4-2.5 Å, for 1UUS and 3IN7 it is 1.3-1.8 Å. One part of BG-loop (e.g. 2K7A – RLRYPVCG) is rigid (the range of Rmsd 1.0-1.8 Å) in comparison with the other part (e.g. 2K7A – YNGGGLVT), which is more flexible (1-4 Å). Also, in BC-loop (1O49 – SETTKGA, 1UUS – SERNPGQ, 2FCI – KRNEPNSY, 2K7A – DSRTPGTY, 2GE9 – SSKAGK, 2K7A – DSRTPG-

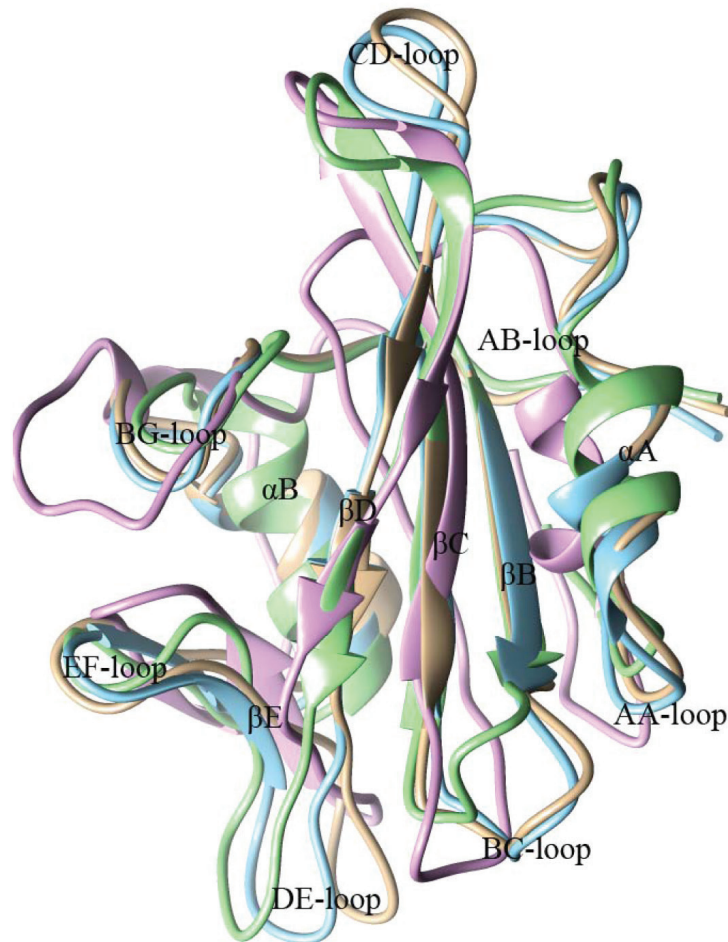


Fig. 1. Example of superposition of different conformation – PDB ID 1O49: grey – the initial form of domain, green – first conformation (RMSD - 3.2 Å); blue – second conformation (2.6 Å); pink – third conformation (5.17 Å)

TY, 2JYQ – SETTKG, 2KK6 – SHGKPGE, 3IN7 – SESAPGD) we observed similar situation. In this case the central part of the loop (e.g. 2FCI – EPN) is more flexible than peripheral parts (e.g. 2FCI – KRN and SY). The RMSD range is 2.0-4.1 Å and 1.5-1.9 Å, respectively. Such mobility could be explained according to the fact that the loop connects the β B- and β C- sheets whose movement is minimal. The flexibility of *CD-loop* (1O49 – NAK, 1UUS – GEVMPAR, 2FCI – KRNEPNSY, 2GE9 – KSTGDPQG, 2K7A – TKAIISENPC, 2JYQ – NAK, 2KK6 – DG, 3IN7 – GN), α B-helix (1O49 – LQQLVAYYSK, 1UUS – FPDFLSE, 2FCI – SLVDLISYYEK, 2GE9 – IPELINYHQ, 2K7A – IPLLIQYHQ, 2JYQ – LQQLVAYY, 2KK6 – IPQLIDHHYT, 3IN7 – LNELVDYHR) and DE-loop, β E-sheet, EF-loop, and FB-loop (1O49 – RKLDSGGFYITSRTQFNS, 1UUS – VQPNDTAAKKT, 2FCI – QQEGQTM-LGNSEFD, 2GE9 – VCSTPQSQYYLAEKHLFST,

2K7A –, 2JYQ – KLDSGGFYITSRTQFSS, 2KK6 – IKETNDSPKRYVVAEKYVFDS, 3IN7 – RDGAGKYFLWVVKFNS) is comparatively the same along the length of secondary structure. The large distinction in flexibility was found only between different structures. So, for 1UUS, 2FCI, 2K7A, 2KK6 and 3IN7 the average flexibility of CD-loop is 1.5 Å, 2.8 Å, 2.7 Å, 1.3 Å and 2.8 Å, respectively. In turn, CD-loop of the structures 1O49, 2GE9 and 2JYQ is much more flexible - 6.1 Å, 3.1 Å and 4.6 Å. The α B-helix mobility of 1UUS, 3IN7 is 1 Å and 1.5 Å. All other structures are more flexible (2.1-2.5 Å).

On the other hand, the flexibility of AA-loop (0.8-2.0 Å), AB-loop (1.4-2.0 Å), β B-sheet (0.9-1.9 Å), β C-sheet (1.2-1.9 Å) and β D-sheet (0.9-2.3 Å) secondary structures is significantly lower. Analysis of *AA-loop* (1O49 – WYFGKIT, 3IN7 – WFFGKIP, 2FCI – WHASLT, 2GE9 – WYSKHMT, 2JYQ – WYFGKIT, 2K7A – WYNKISS, 2KK6 – WYNK-

SIS, 1UUS – IYGYMG) showed that the mobility of the loop is in the range 0.3-8.0 Å (Fig. 1). Almost in all cases four first amino acids of AA-loop are the most flexible (e.g. 1O49 – WYFG, Rmsd changes from 2 to 8 Å); the other amino acids (e.g. 1O49 – KIT) are usually less flexible (Rmsd changes from 0.3 to 2.0 Å). It is not surprising because two (IT) of those amino acids are located in pTyr-binding part of binding pocket. The movement of *AB-loop* (1O49 – LNAENPRGT, 1UUS – LQN-QDPGT, 2FCI – VPRD, 2GE9 – GKEG, 2K7A – GKEGA, 2JYQ – LNPENPRGT, 2KK6 – GKEGA, 3IN7 – SKQRHDGA), *βB-sheet* (1O49 – FLVRE, 1UUS – FIIRF, 2FCI – GAFLVR, 2GE9 – GFIVRD, 2K7A – FMVR, 2JYQ – FLVRE, 2KK6 – FLVRE, 3IN7 – FLIRE), *βC-sheet* (1O49 – YCLSVSDFD, 1UUS – FGIAYI, 2FCI – ATFSR, 2GE9 – YTVS-VFA, 2K7A – TVSVF, 2JYQ – AYCLSVSDFD, 2KK6 – YVLSVYS, 3IN7 – FSLSVKF) and *βD-sheet* (1O49 – GLNVKHYKI, 1UUS – IKHYL, 2FCI – AISFRAEGKIKHCRV, 2GE9 – VIRHYV, 2K7A – IKHYH, 2JYQ – GLNVKHYKIR, 2KK6 – IKHYH, 3IN7 – DVQHFQV) are the same through all length of secondary structure, except few exceptions. So, AB-loop is flexible in the range 1.6-2.0 Å. However, the movement of LNA (1O49) and LNPENP (2JYQ) are 2-3 Å and 1.5-3.0 Å, correspondingly. In turn, the motifs ENPRGT (1O49) and RGT (2JYQ) are less flexible (0.5-1.8 Å and 1.0-1.7 Å). GL (1O49) and DVQ (3IN7) are connected with CD-loop. This fact can be connected with high mobility in those structures in the range from 2.1 to 3.3 Å.

Movement of SH2 domain binding pocket. Eight motifs and single amino acids are presented in SH2 domains binding pocket environment. They are forming pTyr-binding, central and hydrophobic part of SH2 domain binding pocket. The mean flexibility in binding site is lower than in the whole domain structure. It confirms that the binding pocket is more conserved than the rest of the domain. So, its mobility for different structures are: 1O49 – 2.26 Å, 1UUS – 1.27 Å, 2FCI – 2.27 Å, 2GE9 – 2.29 Å, 2JYQ – 2.12 Å, 2K7A – 2.38 Å, 2KK6 – 2.28 Å and 3IN7 – 1.742 Å. However, not all secondary structures give the same contribution to binding site flexibility. pTyr-binding part (forming by AA-loop, αA-helix, βB-sheet, BC-loop and βC-sheet) and hydrophobic pocket (βB-sheet, BC-loop and βC-sheet and “DE-loop, βE-sheet, EF-loop and FB-loop”) are highly flexible.

Results of conformational search

PDB	RMSD limit, Å	Total number of conformations	RMSD, Å (in comparison with initial domain)		
1O49	0.25	3	3.2		
			2.6		
			5.17		
1O49	0.5	1	1.34		
	0.75	1	1.44		
	0.25	1	1.09		
1UUS	0.5	2	1.15		
			3.37		
	0.75	1	2.8		
3IN7	0.25	2	0.94		
			3.74		
	0.5	1	0.95		
3IN7	0.75	1	2.39		
			0.25	2	2.78
			2.12		
2FCI	0.5	2	3.14		
			2.07		
	0.75	1	2.61		
2GE9	0.25	1	2.67		
			2.49		
	0.5	2	2.26		
2GE9	0.75	2	2.17		
			3.10		
	0.25	1	2.3		
2JYQ	0.5	2	2.54		
			2.33		
	0.75	3	2.79		
2JYQ	0.75	3	2.6		
			2.62		
			2.75		
2K7A	0.25	1	2.75		
			2.11		
	0.5	2	3.21		
2K7A	0.75	3	2.87		
			2.95		
	2.03				
2KK6	0.25	1	2.31		
	0.5	1	2.21		
	0.75	2	2.32		
			3.14		

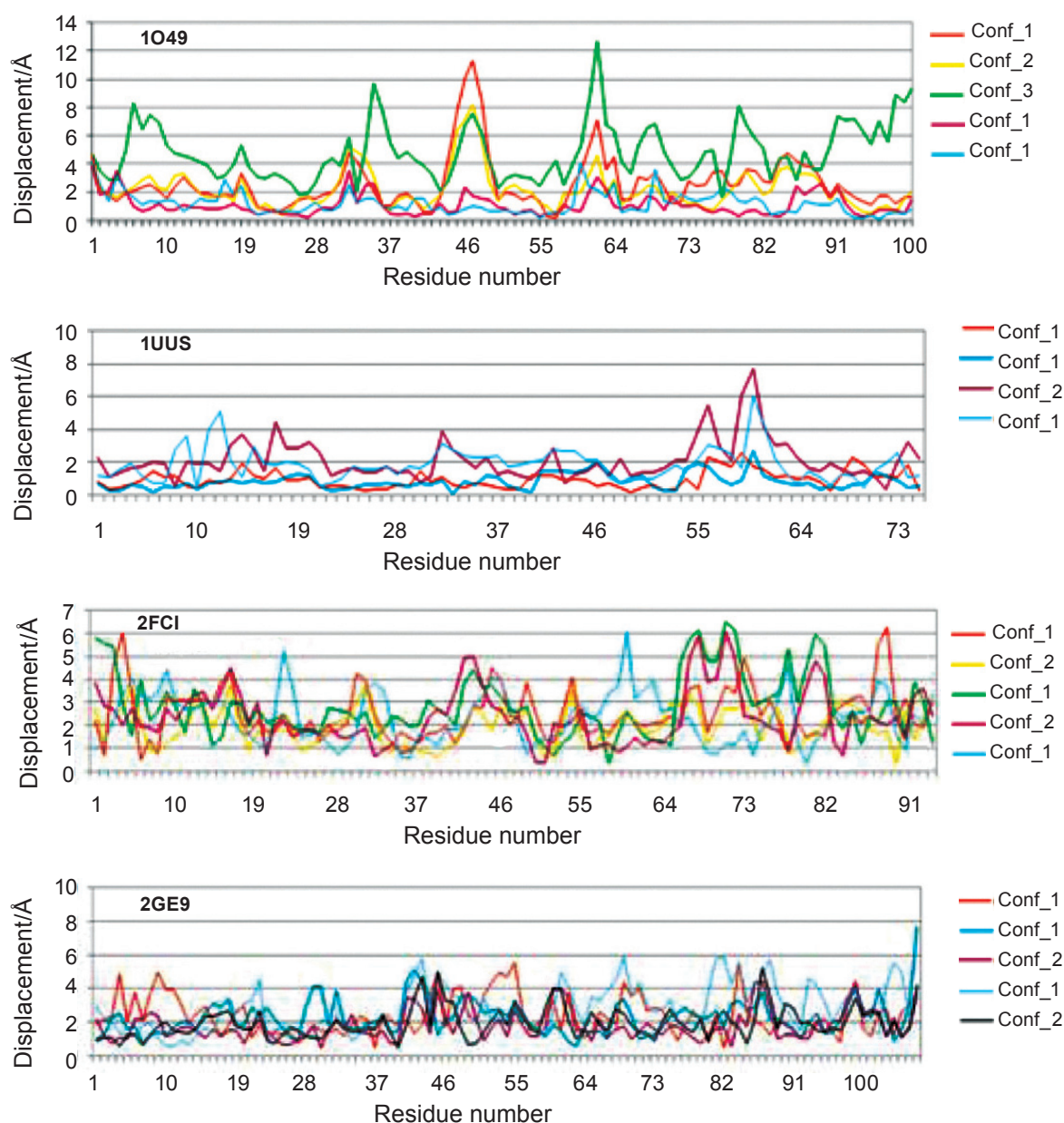


Fig. 2. RMSD of whole conformation represented in Table: (1) lines with RMSD limit 0.25 \AA are marked with: red (conformation N 1); yellow (conformation N 2) and green (conformation N 3); (2) RMSD limit 0.5 \AA : blue (conformation N 1) and (conformation N 2); (3) RMSD limit 0.75 \AA : cyan (conformation N 1); black (conformation N 2) and forest green (conformation N 3)

So, pTyr-binding part is characterized by the movement range $1.14\text{-}2.56 \text{ \AA}$. But, the pTyr-binding part of 1UUS (1.1 \AA) is not characterized by large movement: the movement of pTyr-binding part of 2GE9 (1.7 \AA), 2JYQ (1.8 \AA) and 3IN7 (1.8 \AA) are average; for all the others structures the movement of pTyr-binding part is greater than 2 \AA . Almost whole secondary structures give large contribution to pTyr-binding part mobility: AA-loop (e. g. PDB 1O49 –

2.8 \AA), α A-helix (2.7 \AA), β C-sheet (2.7 \AA), BC-loop (2.8 \AA), except β B-sheet (1.6 \AA).

According to the obtained results, the hydrophobic part of the pocket is more flexible than the other parts. The average movements are: 1O49 – 2.46 \AA , 1UUS – 1.63 \AA , 2FCI – 2.3 \AA , 2GE9 – 2.6 \AA , 2JYQ – 2.3 \AA , 2JYQ – 2.4 \AA , 2K7A – 2.5 \AA , 2KK6 – 2.5 \AA and 3IN7 – 1.7 \AA . As in the previous case, contribution of different secondary structures to

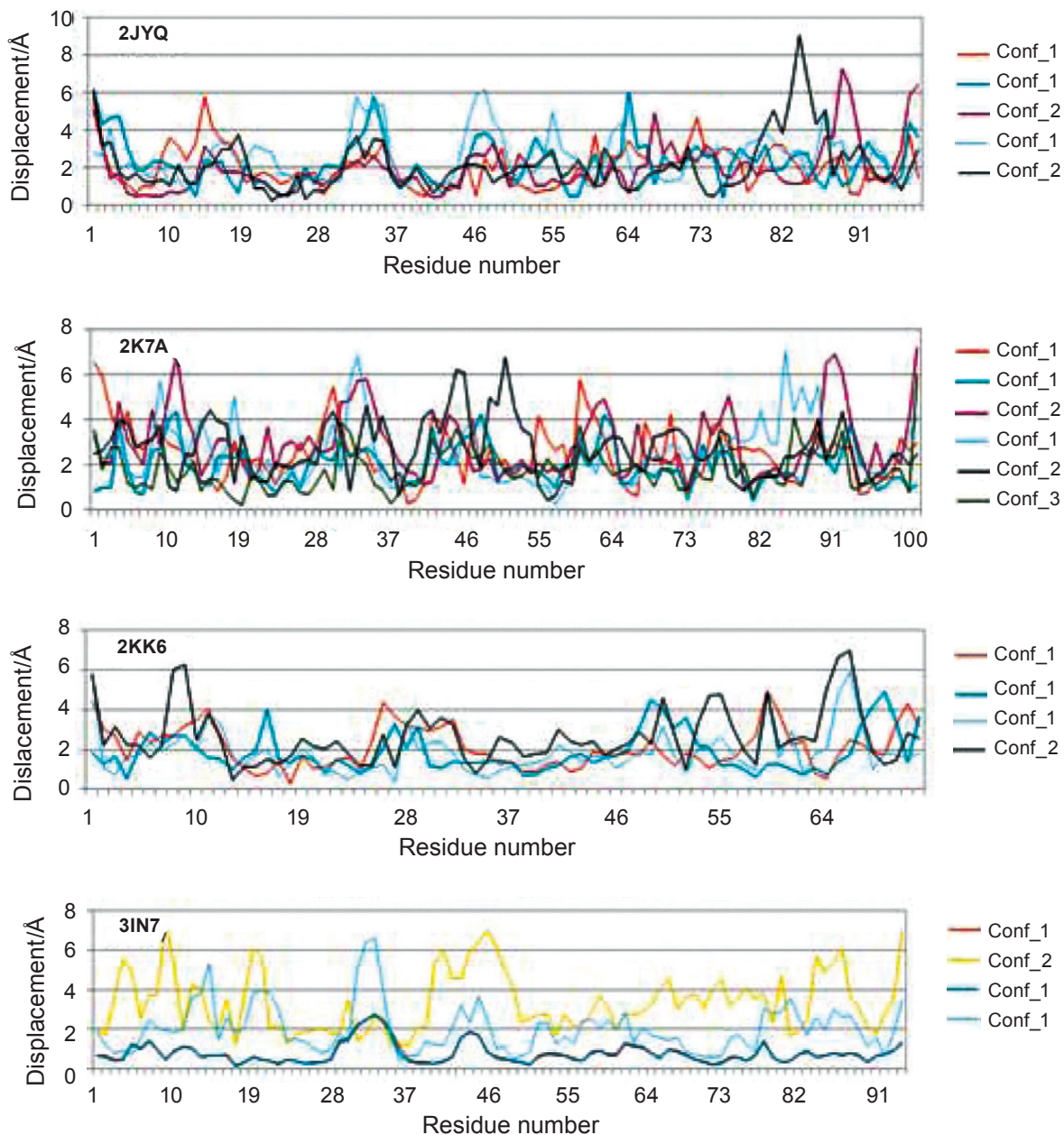


Fig. 2. RMSD of whole conformation represented in Table: (1) lines with RMSD limit 0.25 \AA are marked with: red (conformation N 1); yellow (conformation N 2) and green (conformation N 3); (2) RMSD limit 0.5 \AA : blue (conformation N 1) and (conformation N 2); (3) RMSD limit 0.75 \AA : cyan (conformation N 1); black (conformation N 2) and forest green (conformation N 3)

every unique structure is not the same. For example, in 1O49, 2FCI and 2GE9, BG-loop is highly flexible (2.7 \AA , 2.3 \AA and 3 \AA , respectively) and in 1UUS and 3IN7 one is more rigid (1.3 \AA and 1.7 \AA).

The central part of the pocket is the most stable part. Only in two cases (2FCI – 2.3 \AA and 2GE9 – 2.5 \AA) RMSD is greater than 2 \AA . In all other situations RMSD is less (1O49 – 1.3 \AA , 1UUS – 1 \AA , 2JYQ – 1.8 \AA , 2K7A – 2 \AA , 2KK6 – 1.8 \AA and

3IN7 – 1.6 \AA). Almost in all cases β D-sheet and α A-helix create that part. But there are few exceptions. So, in 1UUS in the creation of this part of pocket BC-sheet (TYR 625) and “DE-loop, β E-sheet, EF-loop and FB-loop” part (VAL 639 and GLN 640) participate too; in 2FCI – CD-loop and in 2K7A – BC-sheet and “DE-loop, β E-sheet, EF-loop and FB-loop”.

It was found that α A-helix, BC-loop, CD-loop, α B-helix, BG-loop and “DE-loop, β E-sheet, EF-loop and FB-loop” secondary structures demonstrate essential contribution to the total variability of SH2 domain conformations. In opposite to this, the flexibility of AA-loop, AB-loop, β B-sheet, β C-sheet, β D-sheet secondary structures are lower. The most flexible conformation was found in PDB file 1O49, namely, conformation N3 (RMSD limit 0.25 Å). The binding pocket is less flexible than other parts of the domain. The hydrophobic and pTyr-binding parts are the most flexible in the binding pocket, but the central part is rigid.

The results may help in further understanding of SH2 domain functioning and provide a starting point for computer simulations of ligand binding.

НОВІ КОНФОРМАЦІЙНІ ВЛАСТИВОСТІ КИШЕНІ ЗВ'ЯЗУВАННЯ ДОМЕНУ SH2

*В. В. Гурмач¹, М. О. Платонов²,
Ю. І. Прилуцький¹*

¹Київський національний університет
імені Тараса Шевченка, Україна;

²Інститут молекулярної біології
і генетики НАН України, Київ;
e-mail: gyrmach@gmail.com

Конформаційні зміни протеїнів відіграють важливу роль в їх біологічному функціонуванні, зокрема взаємодіях типу «ліганд–протеїн» і «протеїн–протеїн». Метою роботи було вивчення конформаційної рухливості найбільш репрезентативних доменів SH2. Виявлено, що кишеня зв'язування домену SH2 включає як гнучкі, так і негнучкі ділянки: найбільш негнучкою є центральна частина, а гнучкими – рТир-зв'язуюча та гідрофобна частини. Результати комп'ютерного аналізу демонструють нові конформаційні властивості доменів SH2, які сприятимуть майбутній розробці лікарських засобів.

Ключові слова: домен SH2, конформаційна рухливість, кишеня зв'язування.

НОВЫЕ КОНФОРМАЦИОННЫЕ СВОЙСТВА КАРМАНА СВЯЗЫВАНИЯ ДОМЕНА SH2

*В. В. Гурмач¹, М. О. Платонов²,
Ю. И. Прилуцкий¹*

¹Киевский национальный университет
имени Тараса Шевченко, Украина;

²Институт молекулярной биологии и
генетики НАН Украины, Киев;
e-mail: gyrmach@gmail.com

Конформационные изменения протеинов играют важную роль в их биологическом функционировании, в частности во взаимодействиях типа «лиганд–протеин» и «протеин–протеин». Целью работы было изучение конформационной подвижности наиболее репрезентативных доменов SH2. Обнаружено, что карман связывания доменов SH2 включает как гибкие, так и негибкие участки: наиболее негибкой является центральная часть, а гибкими – рТир-связывающая и гидрофобная части. Результаты компьютерного анализа демонстрируют новые конформационные свойства доменов SH2, способствующие будущей разработке лекарственных средств.

Ключевые слова: домен SH2, конформационная подвижность, карман связывания.

References

1. Silva C. M. Role of STATs as downstream signal transducers in Src family kinase-mediated tumorigenesis. *Oncogene*. 2004;23(48):8017-8023.
2. Hurmach V. V., Balinskyi O. M., Platonov M. O., Boyko A. N., Borysko P. O., Prylutskyu Yu. I. Design of potentially active ligands for SH2 domains by molecular modeling methods. *Biopolym. Cell*. 2014;30(4):321-325.
3. Hurmach V. V., Platonov M. O., Boyko O. M., Prylutskyu Yu. I. Comparative analysis of SH2 domain structures. *Studia Biologica*. 2015;9:5-14.

4. Hurmach V. V., Balinskyi O. M., Platonov M. O., Boyko O. M., Prylutskyi Yu. I. Application of the methods of molecular modeling to the search for new biologically active substances. *Ukr. Biochem. J.* 2015;87(1):109-20.
5. Pawson T., Gish G. D., Nash P. SH2 domains, interaction modules and cellular wiring. *Trends Cell Biol.* 2001;11(12):504-511.
6. Liu B. A., Jablonowski K., Raina M., Arcé M., Pawson T., Nash P. D. The human and mouse complement resource of SH2 domain proteins—establishing the boundaries of phosphotyrosine signaling. *Mol. Cell.* 2006;22(6):851-868.
7. Andreichenko K. S., Prylutska S. V., Medynska K. O., Bogutska K. I., Nuri-shchenko N. E., Prylutskyi Yu. I., Ritter U., Scharff P. Effect of fullerene C₆₀ on ATPase activity and superprecipitation of skeletal muscle actomyosin. *Ukr. Biokhim. Zhurn.* 2013;85(2):20-26.
8. Corbeil C. R., Williams C. I., Labute P. Variability in docking success rates due to dataset preparation. *J. Comput. Aided. Mol. Des.* 2012;26(6):775-786.

Received 14.07.2015