Published in final edited form as:
*Proteins*. 2010 April ; 78(5): 1326–1330. doi:10.1002/prot.22656.

# NMR Structure of F-Actin Binding Domain of Arg/Abl2 from *Homo sapiens*

**Gaohua Liu**[*], **Rong Xiao**, **Dongyan Wang**, **Yuanpeng J. Huang**, **Thomas B. Acton**, and **Gaetano T. Montelione**[*]
Center for Advanced Biotechnology and Medicine, Department of Molecular biology and Biochemistry, Rutgers, The State University of New Jersey, Piscataway, New Jersey 08854, and Northeast Structural Genomics Consortium

**Keywords**

Abl2; ARG; Abl; Abelson Tyrosine Kinase; Abl Related Gene; helices bundle

## INTRODUCTION

The Northeast Structural Genomics Consortium has used bioinformatics methods to construct a Human Cancer Pathway Interaction Network (HCPIN),1 a comprehensive 3D structure-function database of human-cancer-associated proteins and protein complexes in the context of their interaction networks. The FABD domain of Arg (Abl-related gene; Abl2) was selected as NESG HCPIN target HR5537. Arg and Abl (Abelson tyrosine kinase; Abl1), the Abl non-receptor tyrosine kinases, link diverse cell surface receptors to the regulation of cytoskeletal dynamics and regulate cytoskeletal reorganization, cell proliferation, survival, adhesion, migration and stress responses in multiple cells types.2⁻7 Abl and Arg kinases are multi-domain proteins with highly conserved Src kinase homology 3 (SH3), SH2, kinase (SH1) domains in the N-terminal half. The C-terminal halves of these kinases are more divergent, however, the functions encoded by the C-terminus are critical for the overall functions of these proteins.4 Although Abl and Arg exhibit overlapping expression in many tissues, Arg is most highly expressed in brain, thymus, spleen, and muscle.8 Differences in regulation of cell migration by Abl and Arg have also been reported.7 Consistent with the nuclear and cytoplasmic localization of Abl and the predominant cytoplasmic localization of Arg, three nuclear localization signals (NLS), one nuclear export signal NES motifs and a DNA-binding domain are found in Abl but not in Arg.4 Abl and Arg share a C-terminal calponin homology F-actin-binding domain (FABD) with ~44% sequence identity, which distinguishes Abl family kinases from other non-receptor tyrosine kinases.4،9،10 Preceding this shared FABD, Arg has a microtubule-binding (MT) domain and a second talin-like F-actin-binding domain that is characterized by an I/LWEQ sequence while Abl kinase has a globular (G)-actin binding domain. 4،9،10،11 Arg uses its FABD to anchoring actin and other cytoskeletal partner for signal transfer and other biological functions.4 Both human Abl FABD and Arg FABD belong to the F_actin_bind protein domain family (Pfam12 entry PF08919) comprised of 21 sequences. The NMR structure of the human Abl FABD, the only available structure in F_actin_bind family, has been reported recently13 ; it forms a compact left-handed four helix bundle in solution. The Arg FABD was selected as HCPIN target by NESG for

[*]Correspondence to: Gaohua Liu, CABM, Rutgers University, 679 Hoes Lane, Piscataway, NJ 08854, gliu@cabm.rutgers.edu Phone: (732) 235-5327, Fax: (732) 235-5779 or Gaetano T. Montelione, CABM, Rutgers University, 679 Hoes Lane, Piscataway, NJ 08854, guy@cabm.rutgers.edu, Phone: (732) 235-5321, Fax: (732) 235-5779.

structure determination.1 The NMR structure of human Arg FABD reported here can serve as a structural basis for elucidating the molecular mechanism of Arg pathway, for studies of protein complex formation, and potentially in small molecule drug design.

## MATERIALS AND METHODS

The Arg (Abl2) F-actin binding domain (FABD) from *Homo sapiens* (UniProtKB/Swiss-Prot ID P42684/ABL2_HUMAN, residues 1058–1182) was cloned, expressed and purified following standard, largely automated NESG protocols to produce a uniformly $^{13}C$, $^{15}N$-labeled protein sample.14 Briefly, the truncated ABL2_HUMAN (1058–1182) gene was cloned into a pET14-15C (Novagen) derivative, yielding the plasmid pHR5537A-14.12. The resulting construct contains 10 nonnative residues at the N-terminus (MGHHHHHHSH) to facilitate protein purification and one single mutation T1062A was introduced. *Escherichia coli* BL21 (DE3) pMGK cells were transformed with pHR5537A-14.12, and cultured in MJ9 minimal medium15 containing $(^{15}NH_4)_2SO_4$ and *U* -$^{13}C$-glucose as sole nitrogen and carbon sources. *U*-$^{13}C$, $^{15}N$ Arg FABD was purified using an ÄKTAxpress™ (GE Healthcare) based two step protocols consisting of IMAC (HisTrap HP) and gel filtration (HiLoad 26/60 Superdex 75) chromatography. The final yield of purified *U*-$^{13}C$, $^{15}N$ Arg FABD (> 98% homogeneous by SDS-PAGE; 15.2 kDa by MALDI-TOF mass spectrometry) was ~100 mg/L. In addition, a *U*-$^{15}N$ and 5% biosynthetically directed fractionally $^{13}C$-labeled sample16 was generated for stereo-specific assignment of isopropyl methyl groups. Both *U*-$^{13}C$,$^{15}N$ and 5%$^{13}C$,*U*-$^{15}N$ Arg FABD were dissolved, respectively, at concentrations of ~1.2 mM and 1.4 mM in 95% $H_2O$/5% $D_2O$ (20 mM MES, 200 mM NaCl, 10 mM DTT, 5 mM $CaCl_2$, 0.02% $NaN_3$) at pH 4.5. Analytical gel filtration and static light scattering data14 indicates that this domain is monomeric in solution under the conditions used for these NMR studies.

All NMR spectra were recorded at 25 °C. Triple resonance NMR data (3D HNCO, 3D HNCACB, 3D CBCAcoNH, and (4,3)D GFT $H^{\alpha\beta}C^{\alpha\beta}$coNHN17) were collected on Varian INOVA 600 MHz, a simultaneous 3D $^{15}N/^{13}C^{aliphatic}/^{13}C^{aromatic}$-edited NOESY 18 spectrum (mixing time 100 ms) in H2O and a 3D $^{13}C$-edited NOESY in D2O were acquired on a Bruker AVANCE 800 MHz spectrometer. 2D constant-time [$^{13}C$, $^{1}H$]-HSQC spectra with 28 ms and 42 ms constant-time delays were recorded for the 5% biosynthetically directed fractionally $^{13}C$-labeled sample on a Varian INOVA 600 MHz spectrometer equipped with a cryogenic probe in order to obtain stereo-specific assignments for isopropyl groups of valines and leucines.16 All NMR data were processed using the program NMRPipe 19 and analyzed using the program XEASY.20 Spectra were referenced to external DSS. Resonance assignments were achieved as described previously.21 Sequence specific assignments ($H^N$, $H^\alpha$, N, $C^\alpha$) and $H^\beta/C^\beta$assignments were obtained largely automated with the program AUTOASSIGN.22 These assignments were then used to simulate a NOESY peak list that was used to aid interactive side-chain assignments. The simultaneous 3D $^{15}N/^{13}C^{aliphatic}/^{13}C^{aromatic}$-edited NOESY and CCH-TOCSY spectra were then analyzed manually to obtain nearly complete side-chain assignment. Assignments were obtained for 90% of backbone and side-chain chemical shifts assignable with the NMR experiments listed above (excluding N-terminal $NH_3^+$, Lys $NH_3^+$, Arg $NH_2$, OH of Ser, Thr and Tyr, $^{13}C^\gamma$ of Asp and Asn, $^{13}C^\delta$ of Glu and Gln, and aromatic $^{13}C^\gamma$ shifts). Chemical shifts were deposited in the BioMagResBank on 06/14/2009 with accession code 16349.

Based on chemical shifts, the locations of regular secondary structure elements were identified.23 A NOESY peak list containing expected intra-residue, sequential and α-helical medium range NOE peaks was initially generated and was manually edited by visual inspection of the simultaneous NOESY spectra, and subsequent manual peak picking was pursued to identify remaining, primarily long-range NOEs.21 The programs CYANA 24،25

and AUTOSTRUCTURE26 were used in parallel to automatically assign long-range NOEs. Assignments identically obtained by both programs ('consensus assignments') were retained and established the starting point for iterative cycles of noise/artefact peak removal, peak picking, NOE assignment and structure calculation.21 $^{1}$H - $^{1}$H upper distance limit constraints for structure calculations obtained from both NOESY were summarized in Table I. In addition, backbone dihedral angle constraints were derived from chemical shifts using the program TALOS+27 for residues located in well-defined secondary structure elements (Table 1). The final structure calculation was performed with CYANA 3.0, and the 20 conformers with the lowest target function value were refined in an 'explicit water bath' 28 using the program CNS.29 The coordinates were deposited in the Protein Data Bank on 06/14/2009 (accession code 2KK1).

## RESULTS AND DISCUSSIONS

The solution NMR structure of Arg FABD consists of four α-helices α1–α4 (residues 1086–1099, 1106–1123, 1130–1152, 1165–1181) and shows a typical four-helical up-and-down bundle (Fig. 1a & 1b). Structural statistics are given in Table 1; the resulting ensemble of structures exhibits high structure quality assessment scores. The four helices are packed anti-parallel and connected by short "underhand" loops. The N-terminus region from residues 1058–1070 are flexible and disordered, which is consistent with the disorder prediction analysis30 and hetero nuclear NOE data. Some conformers also show a short one turn helix α1' at residues 1080–1083. Most residues buried in the bundle are hydrophobic. Very few inter-helical salt-bridges are observed and the helix bundle is mainly stabilized by the hydrophobic core. The helical bundles, including the hydrophobic core, are well defined by the extensive inter-helical NOE interactions network observed for both neighbor helices and cross-diagonal helices.

Arg FABD shows 44% sequence identity to Abl FABD. The structures of Arg and Abl FABD (PDB 1zzp) are very similar with DALI31 Z score of 12.1, and backbone RMSD of 1.9 Å and 1.2 Å for all 93 aligned residues and all consensus helical residues, respectively (Fig.1c).13 Residues in the hydrophobic core, important for helical packing, are also highly conserved between Arg and Abl FABDs. On the molecular surface, the residues on N-terminal half of the helix α3 are highly conserved between Arg and Abl FABDs. (Fig1d). These conserved surface residues on the N-terminal half of helix α3 and some adjacent surface residues on helix α2 and α4 have been identified to be responsible for F-actin binding and cytoskeletal association.13

Both Arg and Abl FABDs have similar basic charged surface distribution near these conserved F-actin binding area (Fig 1e). The conserved F-actin binding is a potential target area for drug development aimed at inhibiting Arg/Abl function. These structural similarities suggest Arg FABD may bind to F-actin using a similar mechanism as Abl FABD. There are also some obvious differences between Arg and Abl FABD structures. In Arg FABD, the loop between α1 and α2 is proline rich, and four residues shorter comparing to the same loop in Abl FABD. Therefore, the loop between α1 and α2 in Arg FABD is less flexible than the same loop in Abl FABD. The loop between α3 and α4 in Abl FABD is disordered; however, it is shorter and actually ordered in Arg FABD (supported by the hetero nuclear NOE data, data not shown). These different loop regions are not conserved between Arg and Abl FABDs. The helix α3 in Arg FABD extendes two or three more residues on both N- and C-terminal, longer than the helix α3 in Abl FABD. Furthermore, the titled angle of the C-terminal α3 relative to the other three helices in Arg FABD is larger than the corresponding titled angle in Abl FABD.

The helix bundle architecture is emerging as a common interaction module in a number of proteins involved in cytoskeletal regulation.13 Structure alignment using the program DALI31 reveals significant structural similarity between Arg FABD and several other cytoskeletal protein domains that bind F-actin directly or indirectly, including protein domains with low sequence identity (i.e. most of the DALI 'hits' have < 16% sequence identity, except human Abl FABD). These include vinculin head domain (Fig.1f, PDB code 1zvz with DALI Z score 11.6, RMSD 1.9 Å), alpha-catenin (Fig.1f, PDB code 1dow with DALI Z score 11.5, RMSD 2.5 Å), focal adhesion targeting domain (FAT) of focal adhesion kinase (FAK)32 (Fig.1f, PDB code 1ktm with DALI Z score 11.5, RMSD 2.1 Å), talin-1 (Fig.1f, PDB code 2b0h with DALI Z score 10.6, RMSD 2.3 Å). The activated Abl kinases link diverse cell surface receptors to reorganization of the actin cytoskeleton and regulation of chemotaxis, migration, and invasion.4 Arg regulate focal adhesion dynamics, and Arg requires both its kinase activity and its cytoskeleton-binding C-terminal half to fully inhibit focal adhesions.5,7 Similar to the role of FABD in Abl tyrosine kinase, FAT domain is responsible for FAK's localization.32 The structural similarity of Arg C-terminal FABD and the FAK C-terminal FAT domain indicates a possibility that FABD may compete with FAT domain for binding partners to inhibit focal adhesion.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

We thank Dr. James Aramini for helpful discussions. This work was supported by a grant from the National Institute of General Medical Sciences Protein Structure Initiative (U54-GM074958).

## References

1. Huang YJ, et al. Targeting the human cancer pathway protein interaction network by structural genomics. Mol Cell Proteomics. 2008; 7(10):2048–2060. [PubMed: 18487680]
2. Pendergast AM. The Abl family kinases: mechanisms of regulation and signaling. Adv Cancer Res. 2002; 85:51–100. [PubMed: 12374288]
3. Hantschel O, Superti-Furga G. Regulation of the c-Abl and Bcr-Abl tyrosine kinases. Nat Rev Mol Cell Biol. 2004; 5(1):33–44. [PubMed: 14708008]
4. Gu JJ, Ryu JR, Pendergast AM. Abl tyrosine kinases in T-cell signaling. Immunol Rev. 2009; 228(1):170–183. [PubMed: 19290927]
5. Zandy NL, Playford M, Pendergast AM. Abl tyrosine kinases regulate cell-cell adhesion through Rho GTPases. Proc Natl Acad Sci U S A. 2007; 104(45):17686–17691. [PubMed: 17965237]
6. Zandy NL, Pendergast AM. Abl tyrosine kinases modulate cadherin-dependent adhesion upstream and downstream of Rho family GTPases. Cell Cycle. 2008; 7(4):444–448. [PubMed: 18235247]
7. Peacock JG, et al. The Abl-related gene tyrosine kinase acts through p190RhoGAP to inhibit actomyosin contractility and regulate focal adhesion dynamics upon adhesion to fibronectin. Mol Biol Cell. 2007; 18(10):3860–3872. [PubMed: 17652459]
8. Koleske AJ, et al. Essential roles for the Abl and Arg tyrosine kinases in neurulation. Neuron. 1998; 21(6):1259–1272. [PubMed: 9883720]
9. Van Etten RA, et al. The COOH terminus of the c-Abl tyrosine kinase contains distinct F- and G-actin binding domains with bundling activity. J Cell Biol. 1994; 124(3):325–340. [PubMed: 8294516]
10. Wang Y, Miller AL, Mooseker MS, Koleske AJ. The Abl-related gene (Arg) nonreceptor tyrosine kinase uses two F-actin-binding domains to bundle F-actin. Proc Natl Acad Sci U S A. 2001; 98(26):14865–14870. [PubMed: 11752434]

11. Miller AL, Wang Y, Mooseker MS, Koleske AJ. The Abl-related gene (Arg) requires its F-actin-microtubule cross-linking activity to regulate lamellipodial dynamics during fibroblast adhesion. J Cell Biol. 2004; 165(3):407–419. [PubMed: 15138293]
12. Finn RD, et al. Pfam: clans, web tools and services. Nucleic Acids Res. 2006; 34(Database issue):D247–D251. [PubMed: 16381856]
13. Hantschel O, et al. Structural basis for the cytoskeletal association of Bcr-Abl/c-Abl. Mol Cell. 2005; 19(4):461–473. [PubMed: 16109371]
14. Acton TB, et al. Robotic cloning and Protein Production Platform of the Northeast Structural Genomics Consortium. Method Enzymol. 2005; 394:210–243.
15. Jansson M, et al. High-level production of uniformly N-15- and C-13-enriched fusion proteins in Escherichia coli. Journal of Biomolecular Nmr. 1996; 7(2):131–141. [PubMed: 8616269]
16. Neri D, Szyperski T, Otting G, Senn H, Wuthrich K. Stereospecific nuclear magnetic resonance assignments of the methyl groups of valine and leucine in the DNA-binding domain of the 434 repressor by biosynthetically directed fractional 13C labeling. Biochemistry. 1989; 28(19):7510–7516. [PubMed: 2692701]
17. Atreya HS, Szyperski T. G-matrix Fourier transform NMR spectroscopy for complete protein resonance assignment. Proceedings of the National Academy of Sciences of the United States of America. 2004; 101(26):9642–9647. [PubMed: 15210958]
18. Shen Y, Atreya HS, Liu GH, Szyperski T. G-matrix Fourier transform NOESY-based protocol for high-quality protein structure determination. Journal of the American Chemical Society. 2005; 127(25):9085–9099. [PubMed: 15969587]
19. Delaglio F, et al. Nmrpipe - a Multidimensional Spectral Processing System Based on Unix Pipes. Journal of Biomolecular Nmr. 1995; 6(3):277–293. [PubMed: 8520220]
20. Bartels C, Xia TH, Billeter M, Guntert P, Wuthrich K. The Program Xeasy for Computer-Supported Nmr Spectral-Analysis of Biological Macromolecules. Journal of Biomolecular Nmr. 1995; 6(1):1–10.
21. Liu GH, et al. NMR data collection and analysis protocol for high-throughput protein structure determination. Proceedings of the National Academy of Sciences of the United States of America. 2005; 102(30):10487–10492. [PubMed: 16027363]
22. Moseley HNB, Monleon D, Montelione GT. Automatic determination of protein backbone resonance assignments from triple resonance nuclear magnetic resonance data. Nuclear Magnetic Resonance of Biological Macromolecules, Pt B. 2001; 339:91–108.
23. Wishart DS, Sykes BD. The C-13 Chemical-Shift Index - a Simple Method for the Identification of Protein Secondary Structure Using C-13 Chemical-Shift Data. Journal of Biomolecular Nmr. 1994; 4(2):171–180. [PubMed: 8019132]
24. Guntert P, Mumenthaler C, Wuthrich K. Torsion angle dynamics for NMR structure calculation with the new program DYANA. Journal of Molecular Biology. 1997; 273(1):283–298. [PubMed: 9367762]
25. Herrmann T, Guntert P, Wuthrich K. Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. Journal of Molecular Biology. 2002; 319(1):209–227. [PubMed: 12051947]
26. Huang YJ, et al. An integrated platform for automated analysis of protein NMR structures. Method Enzymol. 2005; 394:111–141.
27. Shen Y, Delaglio F, Cornilescu G, Bax A. TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. J Biomol NMR. 2009; 44(4):213–223. [PubMed: 19548092]
28. Linge JP, Williams MA, Spronk CA, Bonvin AM, Nilges M. Refinement of protein structures in explicit solvent. Proteins. 2003; 50(3):496–506. [PubMed: 12557191]
29. Brunger AT, et al. Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta Crystallographica Section D-Biological Crystallography. 1998; 54:905–921.
30. Huang, YJ.; Montelione, GT. Dis Meta. Available at http://www-nmr.cabm.rutgers.edu/bioinformatic/disorder/

31. Holm L, Sander C. Dali: a network tool for protein structure comparison. Trends Biochem Sci. 1995; 20(11):478–480. [PubMed: 8578593]
32. Liu G, Guibao CD, Zheng J. Structural insight into the mechanisms of targeting and signaling of focal adhesion kinase. Mol Cell Biol. 2002; 22(8):2751–2760. [PubMed: 11909967]
33. Huang YJ, Powers R, Montelione GT. Protein NMR recall, precision, and F-measure scores (RPF scores): Structure quality assessment measures based on information retrieval statistics. Journal of the American Chemical Society. 2005; 127(6):1665–1674. [PubMed: 15701001]
34. Laskowski RA, Macarthur MW, Moss DS, Thornton JM. Procheck - a Program to Check the Stereochemical Quality of Protein Structures. Journal of Applied Crystallography. 1993; 26:283–291.
35. Word JM, Bateman RC, Presley BK, Lovell SC, Richardson DC. Exploring steric constraints on protein mutations using MAGE/PROBE. Protein Science. 2000; 9(11):2251–2259. [PubMed: 11152136]
36. Bhattacharya A, Tejero R, Montelione GT. Evaluating protein structures determined by structural genomics consortia. Proteins. 2007; 66(4):778–795. [PubMed: 17186527]
37. Landau M, Mayrose I, Rosenberg Y, Glaser F, Martz E, Pupko T, Ben-Tal N. ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. Nucleic Acids Res. 2005; 33(Web Server issue):W299–W302. [PubMed: 15980475]
38. Koradi R, Billeter M, Wuthrich K. MOLMOL: A program for display and analysis of macromolecular structures. Journal of Molecular Graphics. 1996; 14(1):51–55. [PubMed: 8744573]
39. DeLano, WL. The pymol manual. San Carlos, CA: DeLano Scientific; 2002.

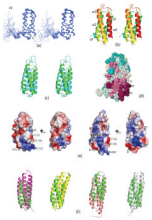

**Fig. 1.**

**Table 1**

Statistics of Arg FABD NMR Structure Determination.

| | |
|---|---|
| Conformationally-restricting distance constraints[a] | |
| Distance constraints | |
| Intra-residue ($i = j$) | 634 |
| Sequential ($\|i– j\| = 1$) | 1017 |
| Medium range ($1 < \|i – j\| < 5$) | 1270 |
| Long range ($\|i – j\| \geq 5$) | 961 |
| Total | 3882 |
| Hydrogen bond constraints[b] | 84 |
| Dihedral angle constraints | 172 |
| φ | 86 |
| ϕ | 86 |
| Number of constraints per residue[c] | 31.3 |
| Number of long range constraints per residue[c] | 7.3 |
| CYANA target function [Å$^2$] | 0.51 ± 0.06 |
| Average r.m.s.d. to the mean CNS coordinates [Å] | |
| regular secondary structure elements [d] backbone heavy atoms N, C$^\alpha$, C’ | 0.36 ± 0.06 |
| regular secondary structure, [d] all heavy atoms | 0.68 ± 0.04 |
| Residues 1078–1182, backbone heavy atoms N, C$^\alpha$, C’ | 0.52 ± 0.09 |
| Residues 1078–1182, all heavy atoms | 0.79 ± 0.08 |
| AutoQF33 R/P/F/DP[e] | 97/94/96/82 |
| Ramachandran plot summary for all ordered (by PSVS) [f] | |
| most favorable regions | 95.0 |
| additionally allowed regions | 4.9 |
| generously allowed regions | 0.0 |
| disallowed regions | 0.0 |
| Global quality scores (raw / Z - score) [f] | |
| VERIFY3D | 0.30 / −2.57 |
| PROSAII | 0.55 / −0.41 |
| PROCHECK G-factor ϕ and φ34 | 0.30 / 1.49 |
| PROCHECK G-factor all dihedral angles | 0.14 / 0.83 |
| MOLPROBITY35 Clash Score | 18.20 / −1.60 |

[a] Analysed for all residues included HisTag .

[b] Helical (i, i+4) hydrogen bond constraints were inferred from structures by using program CYANA, and were applied only in the final refinement.

[c] There are 132 residues with conformationally restricting constraints, first three residues of HisTag was excluded.

[d] Regular secondary element: alpha residues 1086–1099, 1106–1123, 1130–1152, 1165–1181

[e] RPF scores33 (Recall, Precision, F-measurement and DP) reflecting the goodness-of-fit of the final ensemble of structures (including disordered residues) to the NMR data.

[f]Calculated using the protein structure validation software suite (PSVS) 1.3 program, order residues 1065–1066, 1071–1154, 1159–1181, defined based on dihedral angle order parameters S(φ) and S(ψ) > 0.90. Z-scores were computed relative to corresponding structure quality measures for high resolution X-ray crystal structures 36.