W342–W346 *Nucleic Acids Research, 2005, Vol. 33, Web Server issue*
doi:10.1093/nar/gki369

POPSCOMP: an automated interaction analysis of biomolecular complexes

Jens Kleinjung and Franca Fraternali[1,*]

Bioinformatics Unit, Faculty of Sciences, Vrije Universiteit, De Boelelaan 1081A, 1081HV Amsterdam, The Netherlands and [1]Division of Mathematical Biology, National Institute for Medical Research, Mill Hill, London NW7 1AA, UK

Received February 14, 2005; Revised and Accepted March 3, 2005

ABSTRACT

Large-scale analysis of biomolecular complexes reveals the functional network within the cell. Computational methods are required to extract the essential information from the available data. The POPSCOMP server is designed to calculate the interaction surface between all components of a given complex structure consisting of proteins, DNA or RNA molecules. The server returns matrices and graphs of surface area burial that can be used to automatically annotate components and residues that are involved in complex formation, to pinpoint conformational changes and to estimate molecular interaction energies. The analysis can be performed on a per-atom level or alternatively on a per-residue level for low-resolution structures. Here, we present an analysis of ribosomal structures in complex with various antibiotics to exemplify the potential and limitations of automated complex analysis. The POPSCOMP server is accessible at http://ibivu.cs.vu.nl/programs/popscompwww/.

INTRODUCTION

The current focus shift from the analysis of single biomolecules to systems of interacting components requires the development of tools to analyse a multitude of interactions. A key parameter in the interaction of biomolecules is the buried solvent-accessible surface area (SASA) upon complexation, which is readily calculated from the coordinates of a complex structure.

In the early seventies, Lee and Richards (1) defined the solvent-accessible surface as the area described by a probe of the diameter of a water molecule rolling over the protein surface. The calculation of the surface area can be achieved with many methods ranging from more accurate geometric and analytical formulations to discrete approximations, with accuracy being inversely proportional to computational efficiency (2–6). For the analysis of large structural assemblies, a compromise between accuracy and efficiency is needed. Therefore, the SASA calculation chosen for POPSCOMP is a heuristic method that uses a simple analytical formula with a parameterization designed for biomolecules. The formula takes into account single atom areas corrected by multiple overlaps with neighbouring atoms. The details of the formula and the parameterization of the method are described in (7).

Biomolecular interaction surfaces of specific complexes are usually highly complementary, functionally important and therefore well conserved. Analysis of interacting surfaces can help identifying functionally relevant residues in mutation studies or predict interaction sites of potential homologous complex partners. SASA has already been proven to be helpful in the analysis of large protein–protein and protein–nucleic acids complexes (8). Moreover, buried SASA can be related with a change in solvation free energy, yielding an average of 12 and -60 cal/(mol Å^2) for hydrophobic and hydrophilic surface in proteins, respectively (9).

Ribosome structures are among the largest complexes resolved so far to atomic or residue resolution. The ribosome is the core of the protein biosynthesis machinery of the cell. In structural terms, it is an assembly of two subunits, the small 30S subunit and the large 50S subunit; the 30S subunit is composed of 16S RNA and ~20 proteins, while the 50S subunit is composed of 23S RNA, 5S rRNA and ~30 proteins. Ribosomes complexed with antibiotics have revealed insights into the transcription mechanism and its inhibition [reviewed in (10)]. Here, we show the exemplary application of automated complex analysis by extending our previous work on single ribosome structures (11) to selected groups of ribosome structures complexed to various antibiotics.

IMPLEMENTATION

The POPSCOMP server is based on the POPS method, which evaluates the SASA of biomolecules using an analytical formula with parameters that have been optimized on a large

*To whom correspondence should be addressed. Tel: +44 20 8816 2250; Fax: +44 20 8913 8545; Email: ffranca@nimr.mrc.ac.uk

set of diverse structures (7,11). The default method invokes a per-atom parameterization, but low-resolution structures are also tractable through per-residue parameters, where a sphere is centred on each C^{α} (amino acid) or P (nucleotide) atom. The POPSCOMP server splits the specified complex structure into single components. Using the list of complex components, all combinations of pairwise complexes are created. The difference between the SASA of two individual components and the SASA of their pairwise complex yields the buried surface area:

$$\Delta SASA_{buried} = SASA_1 + SASA_2 - SASA_{complex1:2} \quad \mathbf{1}$$

The result is a triangular matrix of $n(n-1)/2$ values of surface burial.

Although computational time is in the order of minutes for standard size protein complexes, ribosome structures take several hours for completion on a standard server machine.

INPUT FORMAT

Protein structures in PDB format can be passed to the server either by specifying their PDB identifier or by uploading a local file. The server usually recognizes chain limits by using the 'TER' delimiter of the PDB format. However, in case the 'TER' delimiters are missing, a text window allows to enter user-specified chain limits as the number of the last atom of each chain. Complex components that are composed of heteroatoms can be included by typing their residue name into the 'HETATM' text window. The button for 'Coarse grained calculation' switches the calculation from the default per-atom analysis to a per-residue analysis, which should be used for low-resolution (C^{α} and P atoms only) structures or for fast calculations on large systems. The button 'Output residue areas' activates the output of SASA per residue, which is the sum of the atomic SASAs for all-atom calculations or the residue SASAs for a coarse-grained calculation. The default output is the total SASA of complex components.

OUTPUT FORMAT

The POPSCOMP server returns the main output as text on the results page. This comprises the hydrophobic, hydrophilic and total SASA of the entire complex, each individual component and all pairwise component complexes (Figure 1) as well as matrices of surface burial in all pairwise complexes (Figure 2). This information is also converted into a graphical output (Figure 3) that can be accessed through links on the results page. Additionally, links are provided to the raw output data of the run.

EXAMPLE ANALYSIS OF RIBOSOME STRUCTURES COMPLEXED WITH ANTIBIOTICS

More than 70 ribosome structures are currently deposited in the PDB structure database (12). The structures selected for this analysis are given in Table 1 together with the source organism, the complexing antibiotics and the literature reference. The top group contains high-resolution 30S structures and the centre group comprises high-resolution 50S

```
========
Total Complex
        Total SASAs of Molecule :
      Total Hydrophobic SASA =     92855.6    Angs^2
      Total Hydrophilic SASA =     90795.2    Angs^2
      Total SASA             =    183650.8    Angs^2
========
Component 1 ; Chain B
        Total SASAs of Molecule :
      Total Hydrophobic SASA =     17834.5    Angs^2
      Total Hydrophilic SASA =     31228.7    Angs^2
      Total SASA             =     49063.2    Angs^2
========
Complex 1:2 ; Chains B:C
        Total SASAs of Molecule :
      Total Hydrophobic SASA =     24835.8    Angs^2
      Total Hydrophilic SASA =     36715.5    Angs^2
      Total SASA             =     61551.3    Angs^2
========
```

Figure 1. Example output listing the surface area of the entire complex, single components (here component 1) and pairwise complexes (here complex 1:2). The original chain enumeration is reported for back-referencing to the PDB structure. Classification into hydrophilic/hydrophobic surface is defined in the atom parameterization, which is accessible through a link on the result page.

```
========
Delta Total SASA
        1      2      3      4      5      6      7
 1            118    299   1558      0      0    440
 2                    28    133      0      0      0
 3                          467      0      0      0
 4                                   0      0    877
 5                                          0      0
 6                                                 0
 7
```

Figure 2. Excerpt of the interaction matrix between all pairwise complexes of the 30S subunit of the ribosome in terms of buried surface area. Matrices for hydrophobic, hydrophilic and total buried area are reported by the POPSCOMP server.

structures. The bottom group contains 50S structures with some low-resolution protein components that were removed before starting the calculations, to avoid mixing of per-atom and per-residue parameters. All three ribosome groups were analysed at per-atom level.

We focus here on two aspects: (i) the direct effects of the complexation with antibiotics and (ii) the overall structural variation between the unaffected parts of the ribosomes in each group, both of which are summarized in Table 2.

SASA effects of antibiotics

Surface burial of antibiotics upon complexing the ribosome are given in columns 3 and 4 of Table 2. Most antibiotics bury between 1/2 and 3/4 of their surface, either exclusively or predominantly in contact with the main RNA components of the ribosome: the 16S (30S subunit) and the 23S (50S subunit). POPSCOMP reports the (known) contacts to proteins (labelled 'S' and 'L') and mRNA. However, the important targets of the antibiotics in Table 2 are the ribosomal RNA sites at or around the catalytically active peptidyl transferase centre or the peptide exit tunnel. These sites are predominantly hydrophobic, matching the surface properties of the antibiotics. Accordingly, the ratio between hydrophobic and hydrophilic SASA contribution to the interaction is about 2/3 hydrophobic to 1/3 hydrophilic for the 30S complexes and 3/4 hydrophobic to 1/4 hydrophilic for the 50S complexes (data not shown).

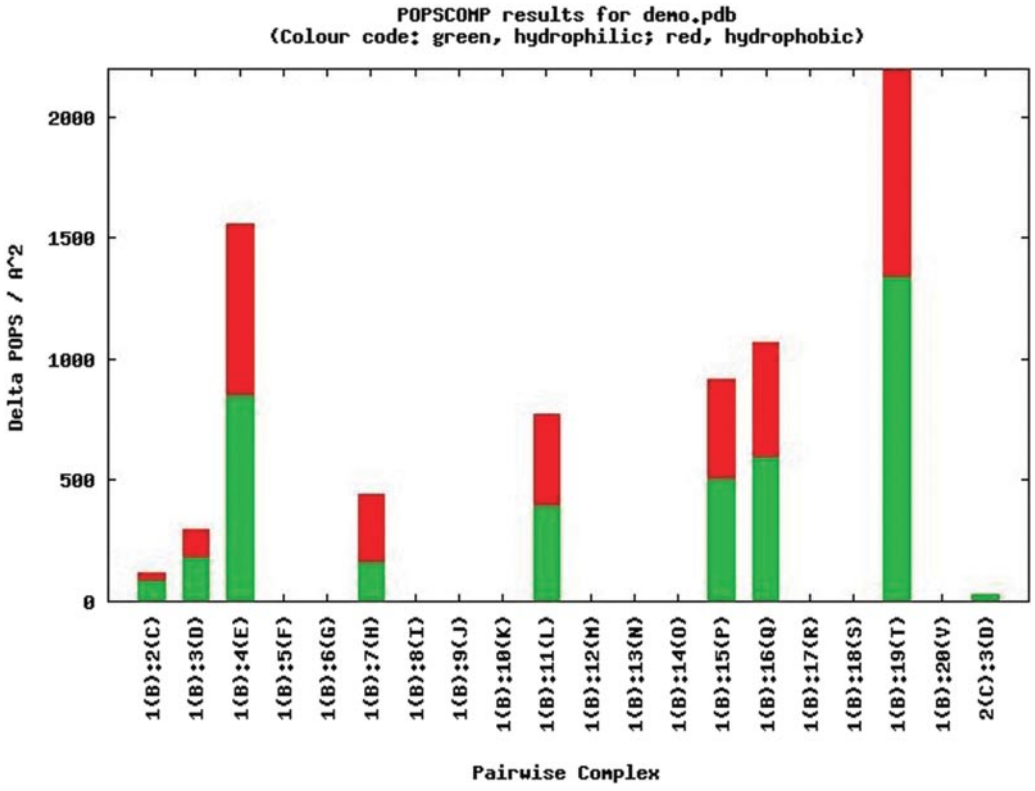

Figure 3. Buried surface areas of all pairwise complexes are given as graphical output. Hydrophobic and hydrophilic SASA contributions are colour coded.

Table 1. Ribosome structures used in automated complex analysis

ID	Ribosome[a]	Antibiotic[b]	Reference
1FJG	30S *T.therm.*	Streptomycin (SRY), spectinomycin (SCM), paromomycin (PAR)	(13)
1HNW	30S *T.therm.*	Tetracycline (TAC)	(14)
1HNX	30S *T.therm.*	Actamycin (PCY)	(14)
1HNZ	30S *T.therm.*	Hygromycin B (HYG)	(14)
1IBK	30S *T.therm.*	Paromomycin (PAR)	(15)
1J5E	30S *T.therm.*	—	(16)
1FFZ	50S *H.maris.*	R(CC)-DA-puromycin (PU)	(17)
1FFK	50S *H.maris.*	—	(18)
1K73	50S *H.maris.*	Anisomycin (ANM)	(19)
1K9M	50S *H.maris.*	Tylosin (TYK)	(20)
1KC8	50S *H.maris.*	Blasticidin S (BLS)	(19)
1KD1	50S *H.maris.*	Spiramycin (SPR)	(20)
1M1K	50S *H.maris.*	Azithromycin (ZIT)	(20)
1N8R	50S *H.maris.*	Virginiamycin M (VIR)	(19)
1NJI	50S *H.maris.*	Chloramphenicol (CLM)	(19)
1Q81	50S *H.maris.*	Puromycin (PPU)	(21)
1Q82	50S *H.maris.*	CC-puromycin (PPU)	(21)
1J5A	50S *D.radio.*	Clarithromycin (CTY)	(22)
1JZX	50S *D.radio.*	Clindamycin (CLY)	(22)
1JZZ	50S *D.radio.*	Roxithromycin (ROX)	(22)
1K01	50S *D.radio.*	Chloramphenicol (CLM)	(22)
1NJM	50S *D.radio.*	ASM/sparsomycin (SPS)	(23)
1NJN	50S *D.radio.*	Sparsomycin (SPS)	(23)
1NJO	50S *D.radio.*	Accpuromycin (PPU)	(23)
1NJP	50S *D.radio.*	t-RNA acceptor stemmimic (PPU)	(23)
1NKW	50S *D.radio.*	—	(24)
1NWX	50S *D.radio.*	abt-773 (773)	(25)
1NWY	50S *D.radio.*	Azithromycin (ZIT)	(25)
1SM1	50S *D.radio.*	Dalfopristin (DOL), quinupristin (SYB)	(26)

[a]*T.therm.*, *Thermus thermophilus*; *H.maris.*, *Haloarcula marismortui*; *D.radio.*, *Deinococcus radiourans.*
[b]Residue names of antibiotics in the PDB structure are given in parentheses.

Table 2. SASA variation of main component and SASA changes of antibiotics upon complexation

ID	$SASA_{molecule\ 1}$ (Å^2)	$SASA_{antibiotic}$ (Å^2)	$SASA_{buried}$ (complex partner) (Å^2)
1FJG	186720	1086 (PAR)	632 (16S), 1 (S12)
		711 (SCM)	553 (16S), 66 (S5)
		978 (SRY)	526 (16S), 164 (S12)
1HNW	191282	1807 (TAC)	741 (16S)
1HNX	187243	881 (PCY)	482 (16S), 69 (mRNA), 29 (S7)
1HNZ	189632	782 (HYG)	324 (16S), 80 (mRNA)
1IBK	188052	1049 (PAR)	585 (16S)
1J5E	188666	—	—
1FFZ	66386	1899 (PU)	776 (23S)
1FFK	331648	—	—
1K73	338985	663 (ANM)	540 (23S)
1K9M	339764	1626 (TYK)	920 (23S), 53 (L22)
1KC8	339261	1903 (BLS)	992 (23S)
1KD1	338273	1589 (SPR)	656 (23S), 1 (L22), 109 (L4)
1M1K	339400	1253 (ZIT)	530 (23S), 2 (L22)
1N8R	339352	914 (VIR)	662 (23S)
1NJI	338361	593 (CLM)	359 (23S)
1Q81	340125	1228 (PPU)	841 (23S), 319 (minihelix)
1Q82	339928	1180 (PPU)	794 (23S), 327 (minihelix)
1J5A	377129	1175 (CTY)	580 (23S)
1JZX	376423	886 (CLY)	593 (23S)
1JZZ	376796	1384 (ROX)	592 (23S)
1K01	375413	617 (CLM)	447 (23S)
1NJM	402458	887 (SPS)	301 (23S)
1NJN	398653	884 (SPS)	296 (23S)
1NJO	398699	930 (PPU)	585 (23S)
1NJP	401691	926 (PPU)	651 (23S)
1NKW	364577	—	—
1NWX	395044	1333 (773)	528 (23S)
1NWY	394805	2063 (ZIT)	551 (23S)
1SM1	414035	1137 (DOL)	715 (23S)
		1554 (SYB)	663 (23S)

Structural variation

The SASAs of the main component of the investigated ribosomes are given in the second column of Table 2. These are taken as the 16S RNA in the 30S subunit and the 23S RNA in the 50S subunit, both molecule '1' in their respective PDB structures. The structural variation that can be expected in large biomolecules at atomic resolution is reflected in the SASA differences between structures 1HNW, 1HNX and 1HNY, which have the same atom composition. The standard deviation between their SASAs is ∼1%. The same holds for the 23S RNA of the second group in Table 2. All structures from 1K73 to 1Q82 have the same composition and the standard variation of SASA is ∼2%.

CONCLUSION

Analysis of biomolecular complexes using the SASA decomposition presented here is a fast and accurate method to obtain information about molecular interactions. Even very large assemblies, such as the ribosome structures, reveal an astonishingly low level of variation of ∼1–2% when analysed in terms of SASA. On the other hand, many antibiotics complexing the ribosome bury ∼50% or more of their surface area. Taken together, these findings suggest that analysis of SASA is a reliable tool for structure analysis and they underline the applicability of automated complex analysis.

ACKNOWLEDGEMENTS

Funding to pay the Open Access publication charges for this article was provided by the National Institute for Medical Research.

Conflict of interest statement. None declared.

REFERENCES

1. Lee,B. and Richards,F.M. (1971) The interpretation of protein structures: estimation of static accessibility. *J. Mol. Biol.*, **55**, 379–400.
2. Richmond,T.J. (1984) Solvent accessible surface area and excluded volume in proteins. Analytical equations for overlapping spheres and implications for the hydrophobic effect. *J. Mol. Biol.*, **178**, 63–89.
3. Wodak,S.J. and Janin,J. (1980) Analytical approximation to the accessible-surface area of proteins. *Proc. Natl Acad. Sci. USA*, **77**, 1736–1740.
4. Hasel,W., Hendrikson,T.F. and Still,W.C. (1988) A rapid approximation to the solvent accessible surface areas of atoms. *Tetrahedron Comput. Methodol.*, **1**, 103–116.
5. Still,W.C., Tempczyk,A., Hawley,R.C. and Hendrickson,T. (1990) Semianalytical treatment of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc*, **112**, 6127–6129.
6. Weiser,J., Shenkin,P.S. and Still,W.C. (1999) Approximate solvent-accessible surface areas from tetrahedrally directed neighbor densities. *Biopolymers*, **50**, 373–380.
7. Cavallo,L., Kleinjung,J. and Fraternali,F. (2003) POPS: a fast algorithm for solvent accessible surface areas at atomic and residue level. *Nucleic Acids Res.*, **31**, 3364–3366.
8. Jones,S., Daley,D.A., Luscombe,N.M., Berman,H.M. and Thornton,J.M. (2001) Protein–RNA interactions: a structural analysis. *Nucleic Acids Res.*, **29**, 943–954.
9. Fraternali,F. and van Gunsteren,W.F. (1996) An efficient mean solvation force model for use in molecular dynamics simulations of proteins in aqueous solution. *J. Mol. Biol.*, **256**, 939–948.
10. Auerbach,T., Bashan,A. and Yonath,A. (2004) Ribosomal antibiotics: structural basis for resistance, synergism and selectivity. *Trends Biotechnol.*, **22**, 570–576.
11. Fraternali,F. and Cavallo,L. (2002) Parameter optimized surfaces (POPS): analysis of key interactions and conformational changes in the ribosome. *Nucleic Acids Res.*, **30**, 2950–2960.
12. Deshpande,N., Addess,K.J., Bluhm,W.F., Merino-Ott,J.C., Townsend-Merino,W., Zhang,Q., Knezevich,C., Xie,L., Chen,L., Feng,Z. *et al.* (2005) The RCSB Protein Data Bank: a redesigned query system and relational database based on the mmCIF schema. *Nucleic Acids Res.*, **33**, D233–D237.
13. Carter,A.P., Clemons,W.M., Brodersen,D.E., Morgan-Warren,R.J., Wimberly,B.T. and Ramakrishnan,V. (2000) Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics. *Nature*, **407**, 340–348.
14. Brodersen,D.E., Clemons,W.M., Carter,A.P., Morgan-Warren,R.J., Wimberly,B.T. and Ramakrishnan,V. (2000) The structural basis for the action of the antibiotics tetracycline, pactamycin, and hygromycin B on the 30S ribosomal subunit. *Cell*, **103**, 1143–1154.
15. Ogle,J.M., Brodersen,D.E., Clemons,W.M., Tarry,M.J., Carter,A.P. and Ramakrishnan,V. (2001) Recognition of cognate transfer RNA by the 30S ribosomal subunit. *Science*, **292**, 897–902.
16. Wimberly,B.T., Brodersen,D.E., Clemons,W.M., Morgan-Warren,R.J., Carter,A.P., Vonrhein,C., Hartsch,T. and Ramakrishnan,V. (2000) Structure of the 30S ribosomal subunit. *Nature*, **407**, 327–339.
17. Nissen,P., Hansen,J., Ban,N., Moore,P.B. and Steitz,T.A. (2000) The structural basis of ribosome activity in peptide bond synthesis. *Science*, **289**, 920–930.
18. Ban,N., Nissen,P., Hansen,J., Moore,P.B. and Steitz,T.A. (2000) The complete atomic structure of the large ribosomal subunit at 2.4 A resolution. *Science*, **289**, 905–920.
19. Hansen,J.L., Moore,P.B. and Steitz,T.A. (2003) Structures of five antibiotics bound at the peptidyl transferase center of the large ribosomal subunit. *J. Mol. Biol.*, **330**, 1061–1075.
20. Hansen,J.L., Ippolito,J.A., Ban,N., Nissen,P., Moore,P.B. and Steitz,T.A. (2002) The structures of four macrolide antibiotics bound to the large ribosomal subunit. *Mol. Cell*, **10**, 117–128.

21. Hansen,J.L., Schmeing,T.M., Moore,P.B. and Steitz,T.A. (2002) Structural insights into peptide bond formation. *Proc. Natl Acad. Sci. USA*, **99**, 11670–11675.
22. Schlünzen,F., Zarivach,R., Harms,J., Bashan,A., Tocilj,A., Albrecht,R., Yonath,A. and Franceschi,F. (2001) Structural basis for the interaction of antibiotics with the peptidyl transferase centre in eubacteria. *Nature*, **413**, 814–821.
23. Bashan,A., Agmon,I., Zarivach,R., Schlünzen,F., Harms,J., Berisio,R., Bartels,H., Franceschi,F., Auerbach,T., Hansen,H.A.S., Kossoy,E., Kessler,M. and Yonath,A. (2003) Structural basis of the ribosomal machinery for peptide bond formation, translocation, and nascent chain progression. *Mol. Cell*, **11**, 91–102.
24. Harms,J., Schlünzen,F., Zarivach,R., Bashan,A., Gat,S., Agmon,I., Bartels,H., Franceschi,F. and Yonath,A. (2001) High resolution structure of the large ribosomal subunit from a mesophilic eubacterium. *Cell*, **107**, 679–688.
25. Schlünzen,F., Harms,J.M., Franceschi,F., Hansen,H.A.S., Bartels,H., Zarivach,R. and Yonath,A. (2003) Structural basis for the antibiotic activity of ketolides and azalides. *Structure*, **11**, 329–338.
26. Harms,J.M., Schlünzen,F., Fucini,P., Bartels,H. and Yonath,A. (2004) Alterations at the peptidyl transferase centre of the ribosome induced by the synergistic action of the streptogramins dalfopristin and quinupristin. *BMC Biol.*, **2**, 4.